SECOND EDITION

Developing Solid Oral Dosage Forms

Pharmaceutical Theory & Practice

Edited by **Yihong Qiu, Yisheng Chen, Geoff G. Z. Zhang, Lawrence Yu, and Rao V. Mantri**

DEVELOPING SOLID ORAL DOSAGE FORMS

This page intentionally left blank

DEVELOPING SOLID ORAL DOSAGE FORMS

Pharmaceutical Theory & Practice

SECOND EDITION

Edited by

YIHONG QIU
Oral Drug Products, Manufacturing Science and Technology, AbbVie, Inc., North Chicago, IL, United States

YISHENG CHEN
Novast Laboratories, Ltd., Nantong, China

GEOFF G. Z. ZHANG
Drug Product Development, Research and Development, AbbVie, Inc., North Chicago, IL, United States

LAWRENCE YU
Office of Pharmaceutical Quality, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD, United States

RAO V. MANTRI
Drug Product Science & Technology, Bristol-Myers Squibb, New Brunswick, NJ, United States

ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Academic Press is an imprint of Elsevier

Academic Press is an imprint of Elsevier
125 London Wall, London EC2Y 5AS, UK
525 B Street, Suite 1800, San Diego, CA 92101-4495, United States
50 Hampshire Street, 5th Floor, Cambridge, MA 02139, United States
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom

Notices
Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

ISBN: 978-0-12-802447-8

Publisher: Mica Haley
Acquisitions Editor: Kristine Jones
Editorial Project Manager: Molly McLaughlin
Production Project Manager: Chris Wortley
Designer: Maria Inês Cruz

Typeset by MPS Limited, Chennai, India

Dedication

This work is dedicated to Drs. R.D. Schoenwald, J. Keith Guillory, L.E. Matheson, the late D.J.W. Grant, G. Amidon, V. Stella, E.L. Parrott, D.R. Flanagan, D.E. Wurster, and P. Veng-Pedersen.

By generously sharing their experience, time, and wisdom, they have taught us well beyond what we learned in school.

We are also forever indebted to our families for their love, understanding, and sacrifice.

This page intentionally left blank

Contents

IV

REGULATORY ASPECTS OF PRODUCT DEVELOPMENT

List of Contributors

G.E. Amidon University of Michigan, Ann Arbor, MI, United States

S. Badawy Bristol-Myers Squibb, New Brunswick, NJ, United States

T. Baxter Jenike & Johanson, Inc., Tyngsboro, MA, United States

J. Brady Ashland Inc., Wilmington, DE, United States

G. Carr Patheon Inc., Mississauga, ON, Canada

S.T. Charlton Bristol-Myers Squibb, Moreton, Wirral, United Kingdom

S. Chatterjee U.S. Food and Drug Administration, Silver Spring, MD, United States

B. Chen AbbVie Inc., North Chicago, IL, United States

W. Chen Gilead Sciences, Foster City, CA, United States

Y. Chen Novast Laboratories, Ltd., Nantong, China

S.C. Das University of Otago, Dunedin, New Zealand

J.Z. Duan U.S. Food and Drug Administration, Silver Spring, MD, United States

T. Dürig Ashland Inc., Wilmington, DE, United States

X.(F.) Fang Novast Laboratories Ltd, Nantong, China

D.R. Flanagan University of Iowa, Iowa City, IA, United States

R.C. Freeze Regulatory Affairs Abbott Laboratories, Abbott Park, IL, United States

Y. Gao AbbVie Inc., North Chicago, IL, United States

Z. Gao U.S. Food and Drug Administration, Silver Spring, MD, United States

C. Gesenberg Bristol-Myers Squibb, New York City, NY, United States

E. Godek OWI-Consulting, Inc., Coatesville, PA, United States

M.Y. Gokhale Drug Product Science & Technology, Bristol-Myers Squibb, New Brunswick, NJ, United States

E. Hadjittofis Imperial College London, London, United Kingdom

X. He Boehringer Ingelheim, Ridgefield, CT, United States

J.Y.Y. Heng Imperial College London, London, United Kingdom

S.W. Hoag University of Maryland, Baltimore, MD, United States

M. Hu Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, PR China; University of Houston, Houston, TX, United States

Y. Hu AbbVie, North Chicago, IL, United States

W. Jiang U.S. Food and Drug Administration, Silver Spring, MD, United States

D. Johnson U.S. Food and Drug Administration, Silver Spring, MD, United States

D. Jones Glatt Air Techniques Inc., Ramsey, NJ, United States

D. Law AbbVie Inc., North Chicago, IL, United States

D. LeBlond CMCStats, Wadsworth, IL, United States

P.I. Lee University of Toronto, Toronto, ON, Canada

S. Lee U.S. Food and Drug Administration, Washington, DC, United States

M. Levin Metropolitan Computing Corporation, East Hanover, NJ, United States

B.V. Li U.S. Food and Drug Administration, Silver Spring, MD, United States

J.-X. Li Ferring Pharmaceuticals, Parsippany, NJ, United States

Q. Li Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, PR China

L. Liu Pfizer Inc., Morristown, NJ, United States

Z. Liu Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, PR China

L. Lu Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, PR China

J.A. Maguire U.S. Food and Drug Administration, Silver Spring, MD, United States

R.V. Mantri Drug Product Science & Technology, Bristol-Myers Squibb, New Brunswick, NJ, United States

R.J. McCann AstraZeneca, Mt. Vernon, IN, United States

M. Mehta U.S. Food and Drug Administration, Silver Spring, MD, United States

P.J. Meyer Pfizer, Inc. Portage, MI, United States

M. Morgen Bend Research, a division of Capsugel Dosage Form Solutions, Bend, OR, United States

D.M. Mudie Bend Research, Bend, OR, United States

A.S. Narang Drug Product Science & Technology, Bristol-Myers Squibb, New Brunswick, NJ, United States

D. Natoli Natoli Engineering Company, St. Charles, MO, United States

A. Newman Seventh Street Development Group, Kure Beach, NC, United States

T. O'Connor U.S. Food and Drug Administration, Washington, DC, United States

B.W. Pack Eli Lilly and Company, Indianapolis, IN, United States

P. Pandey Bristol-Myers Squibb, New Brunswick, NJ, United States

D.Y. Peng Shire Pharmaceuticals Plc, Wayne, PA, United States

S. Porter Ashland Inc., Wilmington, DE, United States

W.R. Porter Abbott Laboratories, Abbott Park, IL, United States

J. Prescott Jenike & Johanson, Inc., Tyngsboro, MA, United States

Y. Qiu AbbVie Inc., North Chicago, IL, United States

K.S. Raghavan Drug Product Science & Technology, Bristol-Myers Squibb, New Brunswick, NJ, United States

J.M. Rowe Bristol-Myers Squibb, New Brunswick, NJ, United States

G. Sackett Freund-Vector Corporation, Cedar Rapids, IA, United States

R. Sanghvi Pharmaceutical Sciences Strategy Allergan Plc., Jersey City, NJ, United States

Z.J. Shao Arena Pharmaceuticals Inc., San Diego, CA, United States

J. Shi Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, PR China

D. Skanchy U.S. Food and Drug Administration, Silver Spring, MD, United States

H. Smith U.S. Food and Drug Administration, Silver Spring, MD, United States

S. Stithit Century Pharmaceuticals, Indianapolis, IN, United States

D. Sun University of Michigan, Ann Arbor, MI, United States

L. Tsygan Metropolitan Computing Corporation, East Hanover, NJ, United States

R.S. Uppoor U.S. Food and Drug Administration, Silver Spring, MD, United States

B. Vig Bristol-Myers Squibb Company, New Brunswick, NJ, United States

J. Wang Actavis, Weston, FL, United States

S. Wang Worldwide Research and Development, Pfizer, Cambridge, MA, United States

J. Wu Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, PR China

L. Wu U.S. Food and Drug Administration, Silver Spring, MD, United States

C. Xie Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, PR China

K. Yamamoto Pfizer Global Research and Development, Groton, CT, United States

T. Yan Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, PR China

Y. Yang U.S. Food and Drug Administration, Silver Spring, MD, United States

A. Yu University of Michigan, Ann Arbor, MI, United States

L.X. Yu U.S. Food and Drug Administration, Silver Spring, MD, United States

F. Zhang University of Texas, Austin, TX, United States

G.G.Z. Zhang AbbVie Inc., North Chicago, IL, United States

H. Zhang U.S. Food and Drug Administration, Silver Spring, MD, United States

X. Zhang U.S. Food and Drug Administration, Silver Spring, MD, United States

L. Zhao U.S. Food and Drug Administration, Silver Spring, MD, United States

P. Zhao U.S. Food and Drug Administration, Silver Spring, MD, United States

Y. Zhao U.S. Food and Drug Administration, Silver Spring, MD, United States

H. Zheng Albany College of Pharmacy and Health Sciences, Albany, NY, United States

J.Y. Zheng Ruby Pharma Inc., Duluth, GA, United States

W. Zheng AstraZeneca Pharmaceuticals LP, Waltham, MA, United States

D. Zhou AbbVie Inc., North Chicago, IL, United States

H. Zhu U.S. Food and Drug Administration, Silver Spring, MD, United States

H.(J.) Zhu Pharmaron, Beijing, China

L. Zhu Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, PR China; AbbVie Inc., North Chicago, IL, United States

PART I

THEORIES AND TECHNIQUES IN THE CHARACTERIZATION OF DRUG SUBSTANCES AND EXCIPIENTS

This page intentionally left blank

CHAPTER

1

Solubility of Pharmaceutical Solids

R.V. Mantri[1], *R. Sanghvi*[2] *and H.(J.) Zhu*[3]

[1]Drug Product Science & Technology, Bristol-Myers Squibb, New Brunswick, NJ, United States [2]Pharmaceutical Sciences Strategy Allergan Plc., Jersey City, NJ, United States [3]Pharmaron, Bejing, China

1.1 INTRODUCTION

1.1.1 Implication of solubility in dosage form development

The solubility of a drug is one of its most important physicochemical properties. The determination of drug solubility and ways to alter it, if necessary, are essential components of pharmaceutical development programs. The bioavailability of an orally administered drug depends primarily on its solubility in the gastrointestinal tract and its permeability across cell membranes. This forms the basis for the biopharmaceutical classification system (BCS).[1] Drug molecules are required to be present in a dissolved form in order for them to be transported across biological membranes. Therefore, low aqueous solubility can either delay or limit drug absorption. Knowledge of the solubility of a drug is also important when direct administration into the bloodstream is desired. Injectable formulations usually require a drug to be in solution for administration. In addition, a drug solution is preferred for conducting pharmacological, toxicological, and pharmacokinetic studies during the drug development stage. Thus, poor aqueous solubility not only limits a drug's biological application, but also challenges its pharmaceutical development. As a result, investigation into approaches for solubility enhancement has been a regular element of pharmaceutical research for several decades. The need for such approaches has been rising following the introduction of combinatorial chemistry and high-throughput screening techniques to the drug discovery arena. The advent of these techniques, resulting in a rapid development of libraries of pharmaceutically active compounds, has led to a greater number of highly active compounds. At the same time, it has resulted in the generation of a far higher percentage of extremely lipophilic and poorly water-soluble compounds, adding more challenges to formulation development. It has been reported[2] that more than a third of the compounds registered by Pfizer in the late 1990s had solubilities that were less than 5 μg/mL.

While solubility enhancement remains one of the primary areas of focus during the drug development phase, there are several situations that may require solubility reduction. Sustained release products, taste masking, and enhancement of chemical stability are examples of such situations.

Knowledge of solubility is also useful when developing analytical methods for drugs. Reverse-phase liquid chromatography is one of the most widely used techniques for pharmaceutical separation and analysis. Separation is based on the differential affinity of the solute toward the mobile phase and the stationary phase, which is a direct outcome of its solubility in these phases. The analysis of concentration using ultraviolet (UV) spectroscopy is also performed on drug solutions.

Based on this discussion, it should be clear that solubility plays an important role in several avenues of pharmaceutical research. As a consequence, the determination of solubility remains one of the most commonly conducted experiments for any new compounds. While solubility experiments are sometimes perceived as trivial, accurate determination of solubility is a challenging exercise. A number of experimental variables may affect the solubility results, inducing high degrees of scatter in the data.[3] This builds a strong case for the applicability of tools for the estimation of solubility based on theoretical calculations. While most of these calculation approaches are useful with respect to providing reasonable estimation and saving time, they can never completely replace the experimentally determined values.

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00001-7

This chapter is written with the intent of developing a thorough understanding of the concepts of solubility. The various physicochemical forces and factors that determine the solubility of a solute in a solvent will be discussed in detail. The thermodynamics of solubilization and various theoretical models for its estimation have also been included. Considerable emphasis has been placed on the techniques used for solubility enhancement, along with practically relevant examples. In addition, the various aspects of solubility determination experiments including challenges and strategies to overcome them are discussed.

1.1.2 Basic concepts of solubility and dissolution

A true solution is a homogenous mixture of two or more components on a molecular level. Any sample collected from such a mixture will be representative of the entire bulk. In a two-component system, the component present in the larger proportion is generally referred to as the *solvent*, and the other as the *solute*.

When a solute is placed in contact with a solvent, mixing occurs due to the propensity of all molecules toward randomization, resulting in an increase in overall entropy of the system. The solute molecules start to break away from the surface and pass into the solvent system. The detached solute molecules are free to move randomly throughout the solvent bulk to form a uniform solution. Some of these solute molecules strike the bulk solute surface and are redeposited onto it. Initially, when the concentration of solute molecules is low in the solution, the number of molecules leaving the bulk solute surface is much higher. As the solvent bulk starts becoming saturated with the solute molecules, the redeposition process starts to accelerate. Once sufficient solute molecules have populated the solvent bulk, the rate of molecules leaving becomes equal to the rate of redeposition (dynamic equilibrium). The concentration of the solute in the solvent at which this equilibrium is reached is defined as the *thermodynamic solubility*, and the rate at which the equilibrium is achieved is the *dissolution rate.* Thus, solubility is an equilibrium concept, while dissolution is a kinetic phenomenon. Both are dependent on the experimental conditions, including temperature. The dissolution rate of a solute in a solvent is directly proportional to its solubility, as described by the Noyes-Whitney equation[4,5]:

$$\text{Dissolution Rate} = \frac{dM}{dt} = \frac{DA}{h}(C_s - C_t) \qquad (1.1)$$

where:

dM/dt is the rate of mass transfer.
D is the diffusion coefficient (cm^2/s).
A is the surface area of the drug (cm^2).
h is the static boundary layer (cm).
C_s is the saturation solubility of the drug.
C_t is the concentration of the drug at time (t).

Solubility is expressed in units of concentration, including percentage on a weight or volume basis, mole fraction, molarity, molality, and parts. The U.S. Pharmacopeia and National Formulary defines the term *solubility* as the number of milliliters of solvent required to dissolve 1 g of the solute.

It follows from the previous discussion that the equilibrium solubility of a solute will depend on its relative affinity toward solvent molecules and fellow solute molecules. Thus, the strength of molecular interactions, both inter and intra, affect solubility. While a detailed description of these interactions can be found in any physical chemistry book, they are discussed here briefly.

1.1.2.1 Ionic interactions

Pure ionic interactions occur between two oppositely charged ions. Such interactions are relevant to pharmaceutical salts and ion pairs. An ion can also interact with a polar molecule (ion-dipole) or induce a dipolar character to a nonpolar molecule (ion-induced dipole). When sodium chloride is dissolved in water, the free sodium and chloride ions interact with polar water molecules such that the positive heads of water molecules interact with the chloride ions, while the negative heads of water molecules interact with the sodium ions. By virtue of these interactions, pharmaceutical salts generally have a higher solubility than their free form. The strength of ionic interactions depends on the electrostatic charge density of the interacting ions, as well as the media properties, including dielectric constant and temperature.

1.1.2.2 van der Waals interactions

Two molecules with permanent dipole moments can interact when placed in sufficiently close proximity (dipole-dipole or Keesom interactions). The molecules will try to arrange themselves in a manner that minimizes the energy associated with them. Thus, the positive head of one molecule will position close to the negative head of the other molecule. The positioning, however, may not be ideal due to geometric constraints and random thermal motion of the participating molecules (entropic influence). As a consequence, a situation arises where the participating molecules, on average, spend more time in an aligned position. Strongly polar molecules can induce polar attributes to nonpolar molecules, resulting in dipole-induced dipole (Debye) interactions. The strength of van der Waals interactions is a direct outcome of the dipole moment and polarizability of the participating

molecules, and it is also affected by the media properties, such as temperature.

1.1.2.3 Dispersion interactions

Also known as *London forces*, dispersion interactions occur between any adjacent pair of atoms or molecules when they are present in sufficiently close proximity. These interactions account for the attractive forces between nonionic and nonpolar organic molecules, such as paraffin and many pharmaceutical drugs. The origin of these forces remains unclear, but it is believed that at any given instance, molecules are present in a variety of distinct positions due to thermal oscillations. These positions give rise to a temporary molecular dissymmetry, resulting in a dipole-type characteristic. This instantaneous dipole then induces a polar character to the neighboring molecules and starts interacting with them.

1.1.2.4 Hydrogen bonding

These interactions occur between hydrogen bond–donating groups and strong electronegative atoms such as halogens, oxygen, and nitrogen. Hydrogen atoms become associated with electronegative atoms by virtue of electrostatics and result in the formation of hydrogen bridges. These interactions are prevalent in aqueous and alcoholic systems. A large number of drugs are involved in either intermolecular or intramolecular hydrogen bonding. The aqueous solubility of a drug is directly related to its hydrogen-bonding capability. The higher water solubility of phenol, as compared to benzene and toluene, can be attributed to the former's hydrogen-bonding nature. The evolving field of cocrystals is based on the hydrogen bond interactions between molecules of the drug and the cocrystal. The strength of hydrogen bond interactions depends upon the electronegativity of the participating atoms, as well as the temperature of the media. Since the requirement of ideal positioning is highest for hydrogen-bonding interactions, they are more sensitive to temperature than other interactions.

1.2 THERMODYNAMICS OF SOLUTIONS

In order to grasp the concepts of solubility, it is essential to understand the basic thermodynamics of mixing. This section covers the various thermodynamic aspects that dictate the process of mixing.

1.2.1 Volume of mixing

The volume of mixing, ΔV_{mix}, is the difference between the physical volume occupied by the mixture (V_{uv}) and the sum of physical volumes occupied by the solute (V_u) and solvent (V_v):

$$\Delta V_{mix} = V_{uv} - (V_u + V_v) \tag{1.2}$$

A negative volume of mixing is indicative of strong intermolecular interactions between the solute and solvent molecules. Aqueous solutions of strong electrolytes have significantly large negative volumes of mixing due to the strong hydration of ions in the solution.[3] In the case of most pharmaceutically active compounds, the volume of mixing is small and can be ignored.

1.2.2 Enthalpy of mixing

The enthalpy of mixing, (ΔH_{mix}), is the difference between the sum of enthalpies of the solute (H_u), the solvent (H_v), and of the mixture (H_{uv}):

$$\Delta H_{mix} = H_{uv} - (H_u + H_v) \tag{1.3}$$

From a strictly enthalpic standpoint, mixing is favored if ΔH_{mix} is negative. The excess enthalpy is liberated in the form of heat (ie, an exothermic process).

The energy of mixing (ΔE_{mix}) is related to the enthalpy of mixing as

$$\Delta E_{mix} = \Delta H_{mix} - P \bullet \Delta V_{mix} \tag{1.4}$$

As previously mentioned, ΔV_{mix} is small, so the values of ΔE_{mix} and ΔH_{mix} are close to one another.

1.2.3 Entropy of mixing

The entropy of a pure system is a measure of the randomness of its molecules. Mathematically:

$$S = R \ln \Omega \tag{1.5}$$

where:

Ω is the number of ways that molecules can be present in the system.

Since Ω is always equal to or greater than unity, entropy is either zero or positive. The entropy of a mixture (ΔS_{mix}) is related to the number of ways that the solute and solvent molecules can exist in pure form and in mixtures:

$$\Delta S_{mix} = R \ln \left(\frac{\Omega_{mix}}{\Omega_u + \Omega_v} \right) \tag{1.6}$$

Generally, the molecules have more freedom to move around in a mixture; that is, Ω_{mix} is greater than $\Omega_u + \Omega_v$. As a consequence, ΔS_{mix} is usually positive. However, in rare situations where the molecules have less freedom in the mixture (solute molecules undergoing complexation-like

interactions with solvent molecules),[6,7] ΔS_{mix} can be negative.

The entropy of mixing is related to the composition of the solution as

$$\Delta S_{mix} = -R(X_u \ln X_u + X_v \ln X_v) \quad (1.7)$$

where:

X_u and X_v are the mole fractions of the solute and solvent in the mixture, respectively.

In a two-component system $X_v = 1 - X_u$, ΔS_{mix} can be written as

$$\Delta S_{mix} = -R(X_u \ln X_u + [1 - X_u]\ln[1 - X_u]) \quad (1.8)$$

$$\Delta S_{mix} = -R(X_v \ln X_v + [1 - X_v]\ln[1 - X_v]) \quad (1.9)$$

It follows from Eq. (1.7) that ΔS_{mix} will be highest for solutions containing equimolar amounts of solute and solvent molecules.

The slope of entropy of mixing as a function of solution composition is given by

$$\frac{\partial \Delta S_{mix}}{\partial X_u} = -R \ln X_u + R \ln(1 - X_u) = -R \ln X_v + R \ln(1 - X_v) \quad (1.10)$$

It follows from Eq. (1.10) that the slope will be highest in dilute solutions (when either X_u or X_v is small). Thus, the manifestation of entropy is highest in dilute solution, which explains why thermodynamically, there is some miscibility between all systems.

1.2.4 Free energy of mixing

The free energy of mixing, ΔG_{mix}, determines the possibility and extent of two compounds mixing to form a solution. It combines the effects of enthalpy and entropy on mixing and is mathematically described as

$$\Delta G_{mix} = \Delta H_{mix} - T \bullet \Delta S_{mix} \quad (1.11)$$

where:

T is the temperature in Kelvin.

Like any thermodynamic process, mixing will occur if the free energy of mixing is negative. On the other hand, if the free energy of mixing is greater than zero, there will be phase separation. As mentioned in the previous section, solubility depends on the temperature of the system. It follows from Eq. (1.11) that an increase in temperature will increase the effect of entropy, thus making mixing more favorable. In addition, temperature may affect ΔH_{mix}, particularly for hydrogen-bonded solvents such as water. It is well known that the strength of hydrogen-bonding interactions is very sensitive to temperature. An increase in temperature makes the self-associated structure of water weaker. Since the self-associated structure of water is primarily responsible for the poor aqueous solubility of nonpolar solutes, including drugs, increasing the temperature thereby facilitates their solubility.

1.3 THEORETICAL ESTIMATION OF SOLUBILITY

While solutions of all states of matter (gas, liquid, and solid) exist in practice, the focus of this chapter will be on solutions comprising of liquid solvent since these systems are most commonly encountered and have the highest relevance in the pharmaceutical field. The backbone of the concepts discussed here may be applied to other systems, with some modifications.

1.3.1 Ideal solutions

In an ideal solution, the strength and density of interactions between the solute molecules and the solvent molecules are equal to that in the solution. In other words, the solute-solvent interactions are equal in magnitude to the solute-solute interactions and solvent-solvent interactions. Thus, the solute and solvent molecules have no preference in terms of interacting with other molecules present in the solutions. This results in the enthalpy (ΔH_{mix}^{ideal}) and volume of mixing (ΔV_{mix}^{ideal}) being 0. Such solutions rarely exist in practicality, but understanding the concept of ideal mixing provides a good platform to understand more complex systems. A solution comprising of solute and solvent bearing very close structural resemblance (in terms of functionality and size) may make nearly ideal solutions. A solution of water in heavy water nearly fits the description of an ideal solution.

The entropy of mixing of an ideal solution (ΔS_{mix}^{ideal}) is given by Eq. (1.7). The partial molar entropy of mixing of the solute in a dilute solution is given by

$$\Delta S_{mix}^{ideal}(u) = -R \ln X_u \quad (1.12)$$

Since X_u is always less than 1, ΔS_{mix}^{ideal} is positive. The free energy of mixing for an ideal solution (ΔG_{mix}^{ideal}), which is the difference between the enthalpy and entropy of mixing, will therefore always be negative. Thus, in ideal solutions, a liquid solute will be miscible with the solvent in all proportions:

$$\Delta G_{mix}^{ideal}(u) = TR \ln X_u \quad (1.13)$$

1.3.2 Effect of crystallinity

In the case of an ideal solution of a crystalline solute, more considerations are required. As discussed

previously, the solute molecules have no preference in terms of interacting with other molecules in the solution. However, if the solute exists as a crystalline solid, the enthalpic component related to the crystallinity of the solute also warrants consideration. In other words, the liquid solute molecules are free to move around in the solution, while the crystalline solute molecules have to be removed from the crystal lattice before they can start moving around. Mathematically, this can be stated as

$$X_{\text{solid}}^{\text{ideal}}(u) = X_{\text{liquid}}^{\text{ideal}}(u) - \text{Effect of Crystallinity} \qquad (1.14)$$

The process is conceptually similar to melting of a solid, which is followed by dissolution of the liquid solute molecules, and is governed by the same interactions as melting.

The effect of crystallinity of a solid on its solubility is described by the Clausius-Clapyron equation:

$$R \ln X_{\text{u}}^{\text{ideal}} = \int_{T_{\text{m}}}^{T} \frac{-\Delta H_{\text{m}}^{(at\ T)}}{T^2} dT \qquad (1.15)$$

$X_{\text{u}}^{\text{ideal}}$ represents the ideal solubility of the crystal and the effect of crystallinity on the solubility.

According to Kirchoff's law, the energy of an irreversible process is equal to the energy of a series of reversible processes between the same end points. Therefore, the irreversible enthalpy of melting at any temperature $T(\Delta H_{\text{m}}^{\text{T}})$ can be described as the sum of the enthalpies for the following three reversible processes: heating the solid to its melting point, T_{m} (quantified by the heat capacity of the solid); melting the solid at its melting point (enthalpy of fusion); and cooling the liquid back down to T (quantified by the heat capacity of the liquid). The heat of melting at T can thus be related to its value at the melting point as

$$\Delta H_{\text{m}}^{\text{T}} = \Delta H_{\text{m}}^{\text{Tm}} + C_{\text{p}}^{\text{C}}(T_{\text{m}} - T) - C_{\text{p}}^{\text{L}}(T_{\text{m}} - T) \qquad (1.16)$$

C_{p}^{C} and C_{p}^{L} are the heat capacities of the crystal and liquid form, respectively.

Combining Eqs. (1.15 and 1.16), the effect of crystallinity can be calculated to be

$$R \ln X_{\text{u}}^{\text{ideal}} = -\Delta H_{\text{m}} \frac{(T_{\text{m}} - T)}{T_{\text{m}} T} + \Delta C p_{\text{m}} \left[\frac{T_{\text{m}} - T}{T} - \ln \frac{T_{\text{m}}}{T} \right] \qquad (1.17)$$

where:

$\Delta C p_{\text{m}}$ is the heat capacity of melting.

According to the van't Hoff expression, the heat capacity of melting is close to zero. The Hildebrand[8] expression states that the heat capacity of melting is equal to the entropy of melting $(\Delta H_{\text{m}}/T_{\text{m}})$. This expression has been used by several researchers, including Prausnitz et al.,[9] Grant et al.,[10] and Mishra.[11] Later, Mishra and Yalkowsky[12] compared the mathematical significance of the two expressions and concluded that the results using either expression are close for solids melting below 600 K. Thus, Eq. (1.17) can be simplified to the following form:

$$\log X_{\text{u}}^{\text{ideal}} = -\Delta H_{\text{m}} \frac{(T_{\text{m}} - T)}{2.303 R T_{\text{m}} T} \qquad (1.18)$$

Eq. (1.18) may be further simplified by applying Walden's rule for entropy of melting. Walden[13] showed that the entropy of melting of coal tar derivatives, which can be assumed to represent organic solids like drugs, is constant, at approximately $6.9R$. Martin et al.,[14] Dannenfelser and Yalkowsky,[15] and Jain et al.[16] have successfully verified and extended the applicability of Walden's rule to several organic nonelectrolytes. Applying this approximation, the effect of crystallinity on the solubility of crystalline solute at 25°C can be calculated by

$$\log X_{\text{u}}^{\text{crystalline}} \approx -0.01(MP - 25) \qquad (1.19)$$

where:

MP is the melting point of the solute expressed in °C.

The obvious interpretation of Eq. (1.19) is that the effect of crystallinity will be greater for high-melting solutes. It is intuitive that a high-melting crystalline solid will offer greater resistance toward solubilization. It also follows that an increase in temperature will diminish the effect of crystallinity and therefore result in increased solubility.

For liquid solutes, where there is a complete absence of crystallinity, Eq. (1.19) should be omitted while calculating the solubility. The effect of crystallinity on solubility is demonstrated in Fig. 1.1.

1.3.3 Nonideal solutions

Nearly all solutions encountered in the pharmaceutical arena are not ideal. Nonideal solutions or real solutions are formed when the affinity of solute molecules for each other is different than that toward the solvent molecules, or vice versa. In either case, the enthalpy of mixing is not ideal. This results in a deviation from ideality that is related to the activity coefficient of the solute. The activity of a solute in a solution is equal to the product of its concentration and activity coefficient as

$$\alpha_{\text{u}} = X_{\text{u}} \bullet \gamma_{\text{u}} \qquad (1.20)$$

where:

γ_{u} is the activity coefficient of the solute.

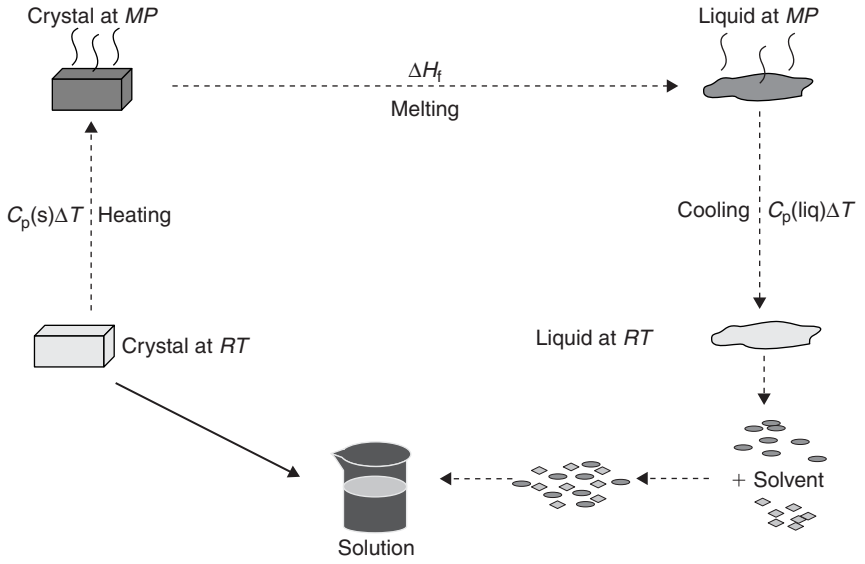


FIGURE 1.1 Role of crystallinity in the solubility of solids. *Source: Modified from a presentation by Dr Kenneth Morris (Short Course on Solubility and Solubilization of Drugs, 2007).*

The mole fraction solubility of a crystalline solute in a nonideal solution can be calculated by combining Eqs. (1.19 and 1.20):

$$\log X_u^{real} = -0.01(MP-25) - \log \gamma_u \qquad (1.21)$$

In the case of ideal solutions, γ_u is 1 and the solubility is ideal.

The mathematical meaning of γ_u is derived using various theories depending upon the type of solute and solvent molecules. Two of the most commonly accepted theories are regular solution theory and aqueous solution theory, discussed next.

1.3.4 Regular solution theory

The regular solution theory is applicable largely to nonhydrogen bonding systems. According to the theory, γ_u is a function of the following three steps, in which mixing occurs:

1. *Removal of a solute molecule from the solute surface.* This involves the breaking of solute-solute bonds, and the work required to accomplish this is given by W_{uu}.
2. *Removal of a solvent molecule from bulk to create a cavity in which the solute molecule can fit.* This involves breaking of solvent-solvent bonds, and the work required for this is given by W_{vv}.
3. *Insertion of the free solute molecule in the created cavity.* This results in gain of work, which is given by W_{uv}. The cavity filling involves the surfaces of both solute and solvent molecules, and that is why the total work gained is $2W_{uv}$.

The total work is therefore equal to $(W_{uu}+W_{vv}-2W_{uv})$. The activity coefficient of a solute is related to this work by the following relationship:

$$\log \gamma_u = (W_{uu}+W_{vv}-2W_{uv}) \bullet \frac{V_u \Phi_v^2}{2.303RT} \qquad (1.22)$$

where:

V_u is the molar volume of the solute.
Φ_v is the volume fraction of the solvent.

It is assumed that W_{uv} is the geometric mean of W_{uu} and W_{vv}. Upon applying this assumption, Eq. (1.22) transforms to

$$\log \gamma_u = (W_{uu}+W_{vv}-2\sqrt{W_{uu}\cdot W_{vv}}) \bullet \frac{V_u \Phi_v^2}{2.303RT} \qquad (1.23)$$

Eq. (1.23) may be rewritten in the following form:

$$\log \gamma_u = (\sqrt{W_{uu}}-\sqrt{W_{vv}})^2 \bullet \frac{V_u \Phi_v^2}{2.303RT} \qquad (1.24)$$

The terms $\sqrt{W_{uu}}$ and $\sqrt{W_{vv}}$ are the solubility parameters of the solute (δ_u) and solvent (δ_v), respectively. Thus, Eq. (1.24) becomes

$$\log \gamma_u = (\delta_u-\delta_v)^2 \bullet \frac{V_u \Phi_v^2}{2.303RT} \qquad (1.25)$$

The solubility parameter is a measure of the cohesive energy density (also referred to as the internal pressure) of a compound:

$$\delta = \sqrt{\frac{\Delta E_v}{V_m}} \qquad (1.26)$$

ΔE_v and V_m are the enthalpy of vaporization and molar volume, respectively.

The solubility parameter is a measure of intermolecular interactions of a compound and can be used as a measure of polarity. Hildebrand and Scott[17] compiled solubility parameters for a number of compounds, and so did Hansen and Beerbower.[18] Hansen and Beerbower further extended the concept by ascribing contributions of the nonpolar, polar, and hydrogen-bonding components of a molecule to the total solubility parameter.

Fedor[19] proposed a scheme to estimate the solubility parameters of liquid organic compounds based on a group contribution approach. Similarly, Sanghvi and Yalkowsky[20] proposed a group contribution scheme to estimate the enthalpy of vaporization of organic compounds.

Eq. (1.21) can be combined with the following equation to calculate the solubility of a crystalline solute according to the regular solution theory:

$$\log X_u^{regular} = -0.01(MP-25)-(\delta_u-\delta_v)^2 \bullet \frac{V_u \Phi_v^2}{2.303RT} \quad (1.27)$$

It follows from Eq. (1.27) that the solubility of a solute is a function of the difference in the cohesive energy densities of the solute and solvent. Thus, a liquid solute (with no crystal term) with a molar volume of 100 mL will be completely miscible in a solvent if their solubility parameter difference is less than 7 $(J/cm^3)^{0.5}$.

The regular solution theory has been applied by several groups to successfully estimate the solubility of organic compounds in nonpolar solvents. It appears that the geometric mean assumption does not seem to hold well for strongly hydrogen-bonded systems, limiting the applicability of the regular solution theory mainly to organic media. Martin et al.[21] proposed a correction factor to cater to the hydrogen-bonding interactions of the solvent. Similarly, the extended Hildebrand solubility approach[8] utilized an extra activity coefficient term to account for strong forces, including hydrogen bonding.

1.3.5 Aqueous solution theory

As discussed in the previous section, the regular solution theory is not applicable to hydrogen-bonding systems like water, since the geometric mean assumption is not valid in such systems. Yalkowsky et al.,[22–24] and later Amidon et al.,[25] proposed that instead of using pressure-volume work to account for enthalpy of mixing, surface tension–area work should be considered. This theory is based on the concept that only the surface of the solute molecule is capable of interacting with solvent molecules. Thus, using surface area instead of volume work is more meaningful. The following equation was proposed to calculate the energy requirement for solubilization in an aqueous system:

$$\Delta E_u^{aqueous} = \Delta H_u^{aqueous} = (\gamma_w + \gamma_u - \gamma_{uw})A_u = \Sigma \gamma_{uw} A_u \quad (1.28)$$

where:

γ_u and γ_w are the surface tensions of the pure solute and water, respectively.
γ_{uw} is the solute-solvent interfacial tension.
A_u is the surface area of the solute.

The entropy of mixing of aqueous systems is also nonideal. The water molecules cluster around an organic solute to form a so-called iceberglike structure. The aqueous solution theory accounts for this deviation by introducing a correction factor in the entropy term:

$$\Delta S_u^{aqueous} = \Delta S_u^{ideal} + \Sigma h_i A_u \quad (1.29)$$

where:

h_i is the contribution of the entropy of mixing per unit area due to the iceberg formation at the molecular surface.

The free energy of mixing the aqueous solution is calculated using Eqs. (1.28 and 1.29):

$$\Delta G_u^{aqueous} = \Sigma \gamma_{uw} A_u - T(\Sigma h_i A_u + \Delta S_u^{ideal}) \quad (1.30)$$

This can be written as

$$\Delta G_u^{aqueous} = \Sigma \gamma_{uw} A_u - T\Sigma h_i A_u + T(RX_u \ln X_u + X_v \ln X_v) \quad (1.31)$$

or

$$\Delta G_u^{aqueous} = \Sigma g_u A_u + T(RX_u \ln X_u + X_v \ln X_v) \quad (1.32)$$

where: $g_u = -\Upsilon_{iw} A_i - T\ (h_i A_i)$ is the group contribution to the activity coefficient, which caters to both the enthalpic and the entropic deviations to ideality.

The solubility of a crystalline solute in a hydrogen-bonding solvent can be calculated by combining Eqs. (1.21 and 1.32):

$$\log X_u^{aqueous} = -0.01(MP-25) - \Sigma g_u A_u \quad (1.33)$$

In a number of studies, Myrdal et al.[26–29] developed a robust group contribution scheme, called the *AQUAFAC approach*, for estimation of the aqueous activity coefficient of organic compounds. Their work successfully demonstrates the applicability of aqueous solution theory.

1.3.6 General solubility equation

The general solubility equation (GSE) utilizes the relation between the octanol-water partition coefficient of a solute and its water solubility, first proposed by Hansch et al.[30]

$$\log S_w = -A \log K_{ow} + B \tag{1.34}$$

where:

K_{ow} is the octanol-water partition coefficient of a liquid solute, while A and B are solute-specific constants.

The octanol-water partition coefficient is the ratio of the activities of the solute in octanol (a_o) and that in water (a_w). For dilute solutions where the activity coefficient is close to unity, K_{ow} can be approximated as

$$K_{ow} = \frac{S_o}{S_w} \tag{1.35}$$

where:

S_o and S_w are the solubilities of the liquid solute in octanol and water, respectively.

Eq. (1.35) can be transformed as follows:

$$\log S_w = \log S_o - \log K_{ow} \tag{1.36}$$

Jain and Yalkowsky[31] proposed that $\log_{10} S_o$ can be replaced by 0.5 for solutes that are completely miscible with octanol. The explanation they provided is as follows.

The concentration of pure octanol is 6.3 M. The saturated solution of a completely miscible solute of a similar molar volume as octanol will contain 3.15 moles of the solute. Since the logarithm of 3.15 is about 0.5, Eq. (1.36) can be modified in the following form to give the molar aqueous solubility of a liquid solute:

$$\log S_w = 0.5 - \log K_{ow} \tag{1.37}$$

It has been discussed previously that, according to the regular solution theory, the miscibility of a solute with a solvent depends on the difference in their solubility parameters and molar volumes. Sepassi and Yalkowsky[32] showed that in the case of octanol (solubility parameter of 21.1 $(J/cm^3)^{0.5}$ and molar volume of 158 cm^3), a liquid solute of molar volume 200 cm^3 will be completely miscible if its solubility parameter is between 15.2 and 27 $(J/cm^3)^{0.5}$. Since the solubility parameters of most drugs are in this range, their liquid forms are expected to be completely miscible in octanol.

Consistent with previous discussions, solute crystallinity should be taken into consideration while calculating the solubility of a crystalline solute. Thus, the solubility of a crystalline solute can be estimated using the GSE by combining Eqs. (1.19 and 1.37):

$$\log S_w = 0.5 - \log K_{ow} - 0.01(MP - 0.5) \tag{1.38}$$

The GSE provides a very simple means of estimating aqueous solubility. Unlike the regular solution theory and aqueous solution theories, which require a number of empirically generated coefficients, the GSE requires only two input parameters (melting point and octanol-water partition coefficient). The measurement of both parameters is a routine part of the active pharmaceutical ingredient (API) characterization process. The value of $\log_{10} K_{ow}$ can also be estimated from the chemical structure by using one of several commercially available software packages (eg, ClogP, ACDlogPdB, and KowWin). Machatha and Yalkowsky[33] compared these estimation schemes and concluded that ClogP was the most accurate predictor of $\log_{10} K_{ow}$.

A variety of schemes for the prediction of melting points of organic compounds have been proposed. These have been reviewed by Katritzky et al.,[34] Dearden,[35] and Bergstrom et al.,[36] among others. Jain and Yalkowsky[37] proposed a reasonably accurate group contribution scheme for melting point estimation using experimental data for over 2200 organic compounds, including a number of drugs.

The accuracy of the GSE has been sufficiently demonstrated for several hundred nonionizable organic compounds.[38–41] Recently,[42] the application of the GSE has been extended to weak ionizable compounds by combining it with the Henderson-Hasselbalch equation.

Notice that in Eqs. (1.37 and 1.38), the solubility is expressed in moles per liter. The constant (0.5) will change to 2.7 ($\log_{10}$500) if the solubility is expressed in grams per liter.

Besides the regular solution theory and aqueous solution theory, several other schemes have been proposed for the estimation of solubility. These were reviewed extensively by Dearden[43] and Jorgensen and Duffy.[44] Abraham and Li[45] developed a model for estimating solubility based on specific solute-solvent interactions. The parameters used in the model include excess molar refraction, dipolarity, polarizability, hydrogen-bond acidity and basicity, and molecular volume of the solute. In addition, hydrogen-bonding terms are included to make the model more sophisticated. This model is very intuitive and simple to follow. However, the coefficients involved with each parameter have to be determined experimentally, and therefore using the model for drug molecules can be very cumbersome. Taskinen and Yliruusi[46] compiled the various prediction schemes based on a neural network approach.

1.4 SOLUBILIZATION OF DRUG CANDIDATES

The solubility of drug candidates can be altered by modifying the crystal form or by changing solvent properties and conditions. As discussed in the previous sections, the equilibrium solubility is dependent

on the properties and nature of interactions in solid state, as well as in solution state. Modifications in the solid state can be accomplished by the use of prodrugs, use of metastable crystal forms, or amorphous state of the drug, salt formation, or co-crystal formation. In solution state, the solubility can be influenced by altering solution pH (for ionizable drugs) or use of additives such as complexing agents, surfactants, or cosolvents. Orally administered drugs must undergo dissolution in the gastrointestinal tract (GI) prior to absorption across the intestinal epithelium. The dissolution rate is not only dependent on the physicochemical properties of the drug, but also the physiological conditions in the GI tract. The in vivo physiological conditions that influence the oral dissolution and solubilization include intestinal pH, GI transit time, bile salts, and lipid content. In the following sections, the impact of solvent pH (ie, drug ionization and salt formation), complexation, micellar solubilization by surfactants, cosolvency, and the combined effects of these factors on drug solubility will be discussed.

1.4.1 Solubility enhancement by pH control and salt formation

A significant portion of drug candidates are weak acids, bases, and their salts.[47,48] As will be shown in the following sections, solubility of these compounds is dependent on the pH of the aqueous medium and the electrolytes in solution. The ionization equilibria and the intrinsic solubility of various unionized (free form) and ionized (salt form) species determine the nature of dependency of drug solubility on pH.

1.4.1.1 Theoretical expressions to describe pH–solubility profiles

The pH-solubility profile of a monoprotic acid (HA) is obtained by the superposition of the two distinct pH-solubility curves for the free form and the salt form, with the constraint that the solubility of either species cannot be exceeded at a given pH (Fig. 1.2). This profile is obtained by following the solubility of free acid (HA) in solutions containing different amounts of a base (eg, MOH).

The profile can be divided into two regions,[48] In Region I, the free acid is present as the excess solid in equilibrium with the solution. The total solubility is the sum of concentrations of unionized (free acid) and the dissociated form:

$$S_T = C_{HA} + C_{A^-} \tag{1.39}$$

At equilibrium (saturation), the concentration of the free acid, C_{HA}, is equal to its intrinsic solubility (S_{HA}).

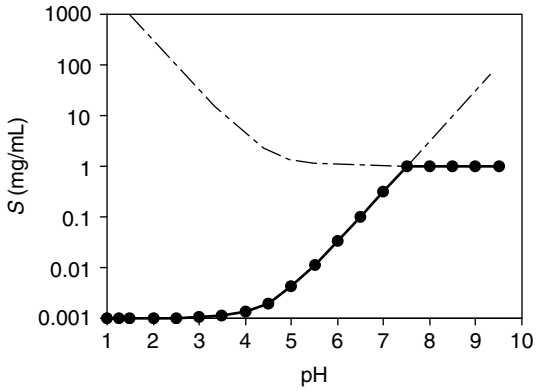


FIGURE 1.2 pH solubility of free acid or its salt.

The concentration of the dissociated form, C_{A^-}, is dependent on the free acid concentration, acid ionization constant, and pH:

$$C_{A^-} = C_{HA}\frac{K_a}{[H^+]} \tag{1.40}$$

Thus, the total solubility of the free acid as a function of pH is given by the following equations:

$$S_T = C_{HA} + C_{A^-} = S_{HA} \bullet \left(1 + \frac{K_a}{[H^+]}\right) \tag{1.41}$$

$$S_T = S_{HA} \bullet (1 + 10^{(pH - pKa)}) \tag{1.42}$$

In Region I, as the pH increases beyond the pK_a of the solute, the concentration of the ionized form increases, leading to an increase in the total solubility in an exponential fashion. This exponential increase in solubility occurs until a point (pH_{max}) where the maximum salt solubility is reached and Region II of the pH-solubility profile begins. In Region II, the salt form is the excess solid in equilibrium with the solution. At equilibrium (saturation), the concentration of the dissociated form (salt), C_{A^-}, is equal to its intrinsic solubility (S_{A^-}). The concentration of the unionized form, C_{HA}, is defined by the ionization equilibrium and is dependent on the ionized form concentration (S_{A^-}) and the acid ionization constant, K_a:

$$S_T = C_{HA} + C_{A^-} = S_{A^-} \bullet \left(1 + \frac{[H^+]}{K_a}\right) \tag{1.43}$$

The solubility of the salt, S_{A^-}, is dependent on the solubility product of the salt, K_{sp}, and the counterion concentration, M^+:

$$K_{sp} = S_{A^-} \bullet [M^+] \tag{1.44}$$

In Region II of Fig. 1.2, the solubility is constant, but this may not be the case when excess counterions (more than stoichiometric amounts) are present in solution, as this will lead to suppression of the salt

solubility (Eq. 1.44). If the counterion M^+ is generated by a weak base, additional ionization equilibria between the charged and uncharged counterion need to be included in Eq. (1.44). The ability of in situ salt formation thus would depend on the pK_a of the acid and pK_a of the conjugate acid of the base.

These equations are expressed in terms of concentration, with the assumption that the activity coefficient for all the species is 1.0. This assumption may be reasonable for neutral species, but for charged species in solutions at high ionic strength, the activity coefficients are less than 1.0. Therefore, the assumption that activity coefficients are unity should be verified.[49] The expressions generated here are valid regardless of whether the excess solid that is added at the beginning of the experiment is in free acid or salt form. The solubility at any given pH, therefore, is governed by pH, the intrinsic solubility of the free acid, the solubility product of the salt, and the concentration of the common ions. The various ionization and solid/liquid equilibria during solubilization of weak acid are shown in Scheme 1.1. It is imperative that the solid phase be analyzed after equilibration during solubility experiments, as the solid phase exists either as the free acid (AH) or its salt (A^-M^+), except at pH_{max}.

By analogy, the pH-solubility profile of a monobasic compound and its salt can be described by the following three equations.

In Region I, when $pH > pH_{max}$:

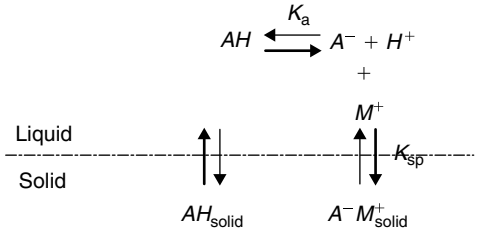
$$S_T = S_B \bullet (1 + 10^{(pKa - pH)}) \tag{1.45}$$

In Region II, when $pH < pH_{max}$:

$$S_T = S_{BH^+} \bullet (1 + 10^{(pH - pKa)}) \tag{1.46}$$

$$K_{sp} = S_{BH^+} \bullet [X^-] \tag{1.47}$$

Fig. 1.3 shows the pH-solubility profile of a free base obtained by titrating with two different acids, *HY* and *HX*. The solubility curves generated in the two experiments (with different acids) show a similar exponential increase in solubility with pH in Region I (freebase is the solid phase). The solubility increases as the pH is lowered until pH_{max}, where the solubility product of the individual salt is limiting.



SCHEME 1.1 Ionization and solid–liquid equilibria. *Source: Drawn based on Serajuddin ATM and Pudipeddi M.* Handbook of pharmaceutical salts. *Wiley-VCH; 2002. p. 135–160 and Chowhan ZT.* J Pharm Sci; *1978;67:1257–60.*

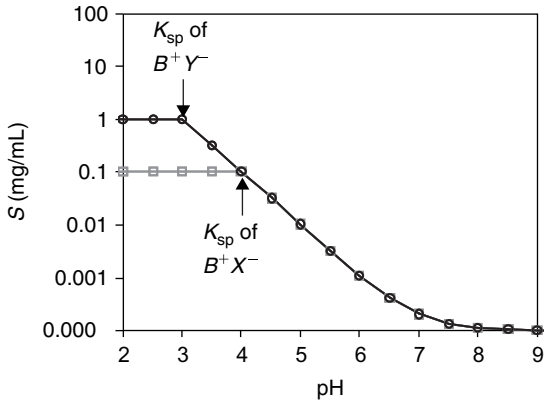


FIGURE 1.3 pH-solubility profile of a free base in the presence of acids *HX* and *HY*.

At equilibrium, at all the pH values (except at pH_{max}), only one solid phase (free base or salt) exists, regardless of the nature of the starting solid material.[51] Therefore, it is important to confirm the identity of the form of the solid phase during solubility experiments. The ratio of the salt solubility and free form solubility drives the location of pH_{max}, which in turn defines the form of the solid phase at equilibrium. The location of pH_{max} is important, as it provides an indication of the relative physical stability of the free form versus salt form in aqueous media and formulations. For example, it may be possible to make a salt of a very weak base (eg, pK_a is approximately 2–3), but when exposed to aqueous conditions in the pH range 1.0–9.0 that are relevant to physiology or formulations, the salt will likely be disproportionate to the free base. A thorough understanding of the pH-solubility profile of ionizable drugs is important for successful development of an ionizable drug candidate.

The discussion in this section has focused on the monoprotic acids and monobasic compounds. However, pH dependency for other weak electrolytes such as zwitterions,[52] dibasic,[48] and diprotic drug candidates can easily be extended by using the same principles of mass balance, ionic equilibria, and the constraint that the solubility of any one of the species cannot be exceeded at a given pH.

1.4.2 Solubilization using complexation

For pharmaceutical systems, *complexation* may be defined as a reversible noncovalent interaction between *m* molecules of drug with *n* molecules of a ligand species[53]:

$$m \cdot D + n \cdot L \overset{K_{m:1}}{\rightleftharpoons} D_m L_n$$

The equilibrium constant $K_{m:n}$ is sometimes known as the complexation or binding or stability constant:

$$K_{m:n} = \frac{[D_m{:}n]}{[D]^m \bullet [L]^n} \tag{1.48}$$

It is possible for a drug to interact with the ligand to form more than one complex, each having a different stoichiometry, and each of these complexation reactions can be defined by the stoichiometry (m:n) and the equilibrium constant ($K_{m:n}$).

In the context of solubility and solubilization, complexation can best be studied using the phase solubility methods described by Higuchi and Connors[54] and Repta.[53] Based on the phase solubility diagram (ie, dependence of solubility on ligand concentration), Higuchi and Connors suggested classifying the complexation phenomena/complexes into the following two categories:

1. Type A phase diagrams, wherein the complex is soluble and does not precipitate regardless of the ligand concentration
2. Type B phase diagrams, wherein the complex precipitates when the ligand concentration reaches a critical value

The type A phase diagrams are further classified into A_L, A_P, and A_N, whereas type B phase diagrams are classified into B_I and B_S phase diagrams. The A_N and B_I types of phase diagrams are neither common nor particularly useful from a solubilization perspective.

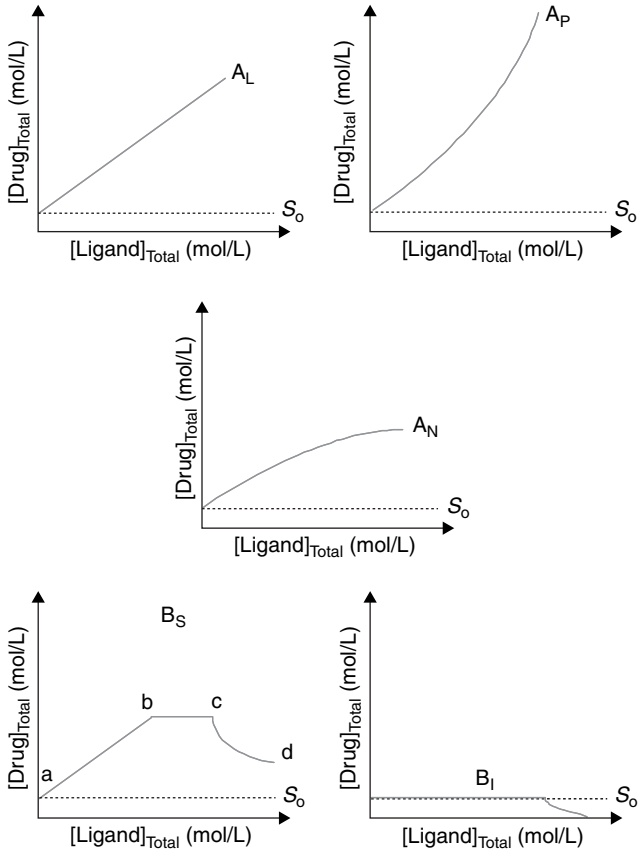


FIGURE 1.4 Drug-ligand phase solubility diagrams. *Source: Classified based on Higuchi T, Connors KA.* Adv Anal Chem; *1965:117.*

1.4.2.1 A_L-type phase diagrams

As shown in Fig. 1.4, A_L-type systems show a linear increase in the drug solubility, S_T, as a function of the ligand concentration, L_T. Such an increase is seen when a soluble drug-ligand complex with a stoichiometry of m:1 is formed. The total drug solubility, S_T is given by

$$S_T = D + m \bullet [D_m L] = S_0 + m \bullet [D_m L] \tag{1.49}$$

By applying mass balance on the ligand concentration and using the definition of equilibrium constant, $K_{m:1}$, the following relationship between total drug solubility and total ligand concentration can be obtained:

$$S_T = S_0 + m \bullet \left(\frac{K_{m:1} \bullet [S_0]^m}{1 + K_{m:1} \bullet [S_0]^m} \right) \bullet [L_T] \tag{1.50}$$

It is noteworthy that an A_L-type phase diagram alone is not sufficient to define the stoichiometry of the soluble drug-ligand complex to be 1:1. As has been demonstrated, total drug solubility linearly increases with total ligand concentration for m:1 drug-ligand complexes.

In recent years, cyclodextrins have received considerable attention as ligands to solubilize hydrophobic drugs by forming inclusion complexes. For an inclusion complex with a stoichiometry of 1:1, the slope of the line from the A_L-type phase diagram can be used to estimate the equilibrium (binding) constant:

$$S_T = S_0 + \left(\frac{K_{1:1} \bullet S_0}{1 + K_{1:1} \bullet S_0} \right) \bullet [CD_T] \tag{1.51}$$

$$K_{1:1} = \left(\frac{\text{Slope}}{(1 - \text{Slope}) \bullet S_0} \right) \tag{1.52}$$

1.4.2.2 A_P-type phase diagrams

The formation of soluble complexes containing more than one molecule of ligand leads to positive deviation from linearity, and such phase diagrams are classified as A_P-type. Let us consider a drug-ligand system forming 1:1 and 1:2 complexes. The complexation equilibria can be written as follows:

$$D + L \overset{K_{1:1}}{\rightleftharpoons} DL$$

$$DL + L \overset{K_{1:2}}{\rightleftharpoons} DL$$

$$S_T = S_0 + K_{1:1} \bullet S_0 \bullet [L] + K_{1:2} \bullet K_{1:1} \bullet S_0 \bullet [L]^2 \tag{1.53}$$

The free ligand concentration, [L], can be related to the total ligand concentration, [L_T], using a mass balance equation for the ligand. Clearly, the total drug solubility will show positive deviation from linearity because, as the concentration of the ligand increases, the contribution of 1:2 complex formation increases. Dilution of drug-ligand systems forming higher-order complexes (A_P-type) may lead to precipitation of the drug, as the concentration of the drug after dilution may be higher than the solubility in the diluted solution.

1.4.2.3 B_S-type phase diagrams

A B_s-type phase diagram can be divided into three regions. In Region I (a–b), the drug solubility initially increases (linearly or nonlinearly, depending on the complex stoichiometry) with an increase in ligand concentration. Region II (b–c), representing the plateau portion, has both the complex and drug in the solid phase, and the total solubility is the sum of intrinsic solubility of the drug and that of the complex. Region III (c–d) begins when the ligand concentration becomes large enough to deplete the free drug concentration, thus decreasing the total drug solubility. At high ligand concentrations, complete removal of the free drug (leading to a solid phase containing pure complex and the total drug concentration) will be equal to the solubility of the complex.

Although drug-ligand solid complexes were reported on in the 1950s,[55,56] they have received considerable attention recently, but by a different name—*cocrystals*.[57–59] Recently, Nehl et al.[60] developed a theoretical framework for cocrystal (drug-ligand complex) phase solubility for cases where the drug and ligand form 1:1 complexes, 1:2 complexes, or both.

The cocrystal solubility (drug concentration in solution with a cocrystal as the solid phase) can be described by the solubility product alone when the two components of the cocrystal do not interact in solution:

$$DL_{solid} \overset{K_{sp}}{\rightleftharpoons} D_{soln} + L_{soln}$$

$$K_{sp} = [D]\bullet[L] \tag{1.54}$$

When the two components in solution also interact to form a 1:1 complex in solution, the total cocrystal solubility is dependent on the solubility product and the equilibrium (binding or complexation) constant:

$$D_{soln} + L_{soln} \overset{K_{1:1}}{\rightleftharpoons} DL_{soln}$$

$$K_{11} = \frac{[DL]}{[D]\bullet[L]} = \frac{[DL]}{K_{sp}} \tag{1.55}$$

The total solubility of the drug is given by

$$S_T = [D]_T = [D]_{free} + [DL] = \frac{K_{sp}}{[L]} + K_{11}\bullet K_{sp} \tag{1.56}$$

The total drug solubility as a function of L_T can be modified to

$$S_T = \frac{K_{sp}}{[L]_T - K_{11}\bullet K_{sp}} + K_{11}\bullet K_{sp} \tag{1.57}$$

This equation describes the total drug solubility as a function of ligand concentration for a phase solubility system that contains a pure, 1:1 drug-ligand cocrystal in the solid phase, and a solution where a 1:1 complex is formed. As predicted by Eq. (1.57), the total solubility decreases as the ligand concentration increases. For the case of the carbamazepine/nicotinamide 1:1 cocrystal, Nehl et al.[60] showed that the total carbamazepine solubility decreased with an increase in nicotinamide concentration, according to Eq. (1.57). In the case where there is no interaction of the drug and ligand in solution (ie, $K_{11} = 0$), the drug solubility is simply defined by the solubility product and ligand concentration in solution.

When studying the phase solubility of complexing agents, it is important to determine the nature of the solid phase along with the drug concentration in the solution phase. Based on the experimentally determined solubility data, a phase diagram illustrating where the drug, ligand, and cocrystal can exist or coexist in the solid phase must be constructed. Understanding the thermodynamic domains of the different forms (drug, ligand, and cocrystal) is essential in the crystallization of the desired cocrystal, preventing undesirable phase transformations during the crystallization and formulation of cocrystals.

1.4.3 Solubilization by cosolvents

The use of cosolvents is well recognized in altering the solubility of organic compounds. Yalkowsky[61] presented some general features of cosolvents and solubilization by cosolvents. Cosolvents are partly polar, due to the presence of hydrogen bond donors, acceptors, or both, thus ensuring miscibility with water. Cosolvents improve the solubility of drugs (nonpolar) because the small hydrocarbon regions of cosolvents reduce the ability of the water to squeeze out nonpolar solutes. The presence of cosolvents in aqueous medium leads to decreased solvent-solvent interactions, leading to a reduction in properties (surface tension, dielectric constant, and solubility parameter) that are reflective of solvent polarity.

The solubility enhancement by a cosolvent is a function of both drug and cosolvent properties.

The greatest enhancement in drug solubility is achieved by cosolvents or cosolvent-water mixtures with similar polarity to the drug. The addition of cosolvent can lead to increase or decrease in solubility, depending on the polarity of the drug candidate. Gould et al.[62] observed that the solubility of oxfenicine is decreased in the presence of ethanol. In some cases, a maximum solubility is observed for an intermediate composition of cosolvent and water, rather than in pure solvents.

Yalkowsky et al.[22,24,63–66] presented a log-linear model to explain the solubility of nonpolar solutes in cosolvent systems. The mixed solvent is assumed to behave similarly to the weighted linear combination of the individual solvents. The general equation describing the solubilization by a cosolvent is defined by

$$\log S_{mix} = \log S_w + \sigma \bullet f_c \tag{1.58}$$

where:

S_{mix} is the solubility in the mixed solvent with f_c as the fractional cosolvent volume, S_w is the solubility in water, and σ is defined as the solubilizing power of the cosolvent.

This equation indicates that the solubility changes exponentially with the cosolvent volume fraction. The logarithm of solubility increases linearly with the volume fraction of the cosolvent, and the slope of this line, σ, is dependent on the properties of both solute and cosolvent:

$$\sigma = S \log K_{ow} + T \tag{1.59}$$

where:

K_{ow} is the octanol-water partition coefficient and is related to the polarity of the drug molecule, whereas S and T are empirically derived cosolvent-dependent constants. The S and T values have been estimated by Millard et al.[66] for a number of cosolvents. The log-linear model is very useful, as it allows the estimation of solubilization of drugs in cosolvent systems, based on the partition coefficient and the known values of S and T, for various pharmaceutically acceptable cosolvents.

Although the cosolvents enhance the solubility of drug candidates by several orders of magnitude, their use is limited due to toxicity, especially at high concentrations. Due to the exponential dependence of drug solubility on volume fraction, dilution may lead to precipitation of the drug, as the drug concentration after dilution may be higher than the solubility in the diluted solution.

1.4.4 Solubilization by surfactants (micellar solubilization)

Surfactants are amphiphilic molecules that self-associate in solution and form aggregated structures known as *micelles*. The two most common models of micellization are the two-phase model and mass action model. From a solubilization perspective, drug molecules interact with micelles to form soluble drug entities. In the two-phase model, the drug is assumed to be incorporated into the micelle, whereas in the mass action model, the drug is assumed to reversibly bind to self-associated surfactant aggregates (micelles).

The two-phase model, also known as the *phase-separation model*, assumes that above a critical concentration of surfactant molecules in aqueous solution, micelles form. The micelles are considered to be a separate phase from the aqueous phase. It is assumed that micellization occurs only above the critical micellar concentration (CMC). Below the CMC, the total surfactant concentration is assumed to be the same as the monomer (aqueous phase) surfactant concentration. Above the CMC, the free monomer (aqueous phase) surfactant concentration is assumed to be constant (equal to the CMC), and the micelle concentration can be obtained by subtracting the CMC from the total surfactant concentration.

Hydrophobic, nonpolar drugs are thought be squeezed out of water into the hydrophobic regions of the micelles.[61] Since the interaction of drug molecules with surfactant monomers is assumed to be negligible, the solubility of the drug in aqueous solution at concentrations below the CMC is assumed to be the same as its intrinsic solubility, S_0. At surfactant concentrations above the CMC the total solubility of the drug, S_T, is given by the following equation:

$$S_T = S_0 + k(P_T - \mathrm{CMC}) \tag{1.60}$$

where:

P_T is the total surfactant concentration drug, and k is a proportionality constant, known as the solubilizing capacity of the surfactant.

The total solubility varies linearly with an increase in total surfactant concentration (ie, micelle concentration). The nonlinearity in solubility enhancement means that the solubilization capacity assumed to be a constant is, in reality, dependent on the surfactant concentration. This can be due to changes in micelle shape or aggregation numbers of the micelle.

Based on the mass action law model, self-association equilibrium between n surfactant monomers and the micelle is expressed as follows:

$$n \cdot P_1 \overset{\beta_n}{\rightleftharpoons} P_n$$

The total surfactant concentration, P_T, can be defined by the following expression:

$$P_T = [P_1] + n \times [P_n] = [P_1] + n \times \beta_n \times [P_1]^n \tag{1.61}$$

Drug solubilization in the surfactant solution can be depicted by multiple equilibria between the self-associated surfactant, P_n, and the drug:

$$i \cdot D + P_n \overset{k_i}{\rightleftharpoons} D_i P_n$$

The saturation solubility of a drug in a surfactant solution is given as

$$S_T = S_0 + [P_n] \sum_{i=1}^{M} i \times k_i \times [S_0]^{in} \tag{1.62}$$

It is difficult to estimate the individual equilibrium constants for the various species, so the two-phase model is a more convenient and often sufficient way to describe the solubility of drugs in surfactant solutions.

1.4.5 Solubilization by combination of approaches

In the previous sections, the most popular techniques to improve solubility by pH control or by use of complexing agents, cosolvents, and surfactants have been discussed. However, it is common to find that a single approach of solubilization is not adequate to improve the aqueous solubility to the desirable extent. Generally, the combination of ionization with cosolvency, complexation, or micellar solubilization leads to synergistic improvement in the solubility of weak electrolytes. The combined effect of complexation and cosolvency, or complexation and micellar solubilization, on drug solubility can be variable. In the following sections, the theoretical framework describing drug solubilization by a combination of approaches is discussed.

1.4.5.1 Combined effect of ionization and cosolvency

The total solubility of an ionizable drug in a mixed solvent can be derived by writing the log-linear model for various drug species in solution. When considering a monoprotic weak acid or monobasic compound, the solubilization by cosolvency for the ionized and unionized drug moieties can be expressed using the log-linear model equations[61]:

$$\log S_u^f = \log S_u + \sigma_u \bullet f_c \tag{1.63}$$

$$\log S_i^f = \log S_i + \sigma_i \bullet f_c \tag{1.64}$$

where:

S_u^f and S_i^f are the solubility of the unionized and ionized species, respectively, in the mixed solvent; S_u and S_i are solubility of the unionized and ionized species, respectively, in water; and σ_u and σ_i are the solubilization power of the cosolvent for the unionized and ionized species. The total drug solubility is obtained by the sum of S_u^f and S_i^f and is provided by the following equation:

$$S_T = S_u \bullet 10^{\sigma_u f_c} + S_i \bullet 10^{\sigma_i f_c} \tag{1.65}$$

For a monoprotic weak acid:

$$S_T = S_0 \bullet 10^{\sigma_u f_c} + S_0 \bullet 10^{(pH - pKa)} \bullet 10^{\sigma_i f_c} \tag{1.66}$$

For a monobasic compound:

$$S_T = S_0 \bullet 10^{\sigma_u f_c} + S_0 \bullet 10^{(pKa - pH)} \bullet 10^{\sigma_i f_c} \tag{1.67}$$

The solubilization capacity of the unionized species, σ_u, is typically found to be greater than that for the ionized species, σ_i, because the cosolvent can solubilize the unionized species (more polar) with greater efficiency than the ionized species (less polar). However, the decrease in solubilization capacity is more than compensated for by the increase in solubility of the ionized species (ie, $S_i >> S_u$). Therefore, it is possible that the combined effect of ionization and cosolvency is better than any single technique. The solubility of 2,2-diphenyl-4-piperidyl dioxolane hydrochloride salt in propylene glycol–water mixtures was better than the solubility of the free base in the mixed solvent or the hydrochloride salt in water.[67]

1.4.5.2 Combined effect of ionization and micellization

The total solubility of a weak electrolyte undergoing ionization and micellization can be described by accounting for the free unionized drug, free ionized drug, micellized unionized drug, and micellized ionized drug:

$$S_T = S_u + S_i + k_u[P_T - CMC] + k_i[P_T - CMC] \tag{1.68}$$

This equation is valid for surfactants that are either neutral or completely ionized in the pH range of interest. Li et al.[68] demonstrated that the falvopiridol (a free base with pK_a of 5.68) showed synergistic improvement in solubility due to the combined effect of lowering pH and using polysorbate 20.

1.4.5.3 Combined effect of ionization and complexation

The combined effect of ionization and complexation on drug solubility has been derived similarly to this scenario of solubilization by ionization and micellization. The total drug solubility has been derived by different researchers[69–71] by considering the various species in solutions (ie, free unionized drug, ionized drug, unionized drug-ligand complex, and ionized drug-ligand complex). Rao and Stella[72] derived an equivalent expression for the total solubility of a weak electrolyte undergoing 1:1 complexes with a ligand, based on the

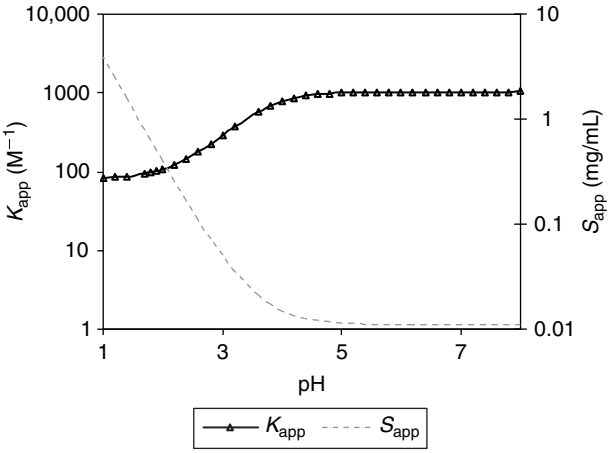


FIGURE 1.5 Variation of S_{app} and K_{app} as a function of pH for thiazolobenzimidazole. *Source: The curves are reproduced based on equations in Yang G, Jain N, Yalkowsky SH.* Int J Pharm*; 2004;269:141–48. and data from Johnson MD, Hoesterey BL, Anderson BD.* J Pharm Sci*; 1994;83:1142–46.*

apparent solubility and apparent binding constant at any given pH, as

$$S_T = S_{app} + \frac{K_{app} S_{app}}{1 + K_{app} S_{app}} [L_T] \tag{1.69}$$

For monoprotic weak acids, K_{app} and S_{app} are defined as follows (Fig. 1.5):

$$K_{app} = K^o \frac{[H^+]}{K_a + [H^+]} + K^- \frac{K_a}{K_a + [H^+]} \tag{1.70}$$

$$S_{app} = S_o \frac{K_a + [H^+]}{[H^+]} \tag{1.71}$$

K_a is the ionization constant; K^o and K^- are the binding constants for the neutral drug-ligand complex and anionic drug-ligand complex, respectively. Similarly, for monobasic compounds,

$$K_{app} = K^o \frac{K_a}{K_a + [H^+]} + K^+ \frac{[H^+]}{K_a + [H^+]} \tag{1.72}$$

$$S_{app} = S_o \frac{K_a + [H^+]}{K_a} \tag{1.73}$$

where:

K_a is the ionization constant; and K^o and K^+ are the binding constants for the neutral drug-ligand complex and cationic drug-ligand complex, respectively.

Rao and Stella[72] presented an interesting analysis of binding constant values for neutral versus ionized species (both acids and bases) with HP- and SBE-β-CD. They reported the ratio of $K^{charged}$ (K^+ or K^-) to K° falls within two orders of magnitude, regardless of the type of cyclodextrins or drugs. However, S_{app} can vary over several orders of magnitude depending on the ionization type (acid or base), its pK_a, and the pH value. Although the variation of K_{app} and S_{app} with pH is in the opposite direction, the magnitude of increase in S_{app} far outweighs the decrease in K_{app} due to ionization, leading to a synergistic increase in total drug solubility, due to ionization and complexation.

1.4.5.4 *Combined effect of cosolvency and complexation*

The combined effects of complexation and cosolvents on drug solubilization can be of a synergistic or antagonistic nature. Some of the factors that need to be considered when using combinations of cosolvent and complexing agents are:

- Due to solubilization by cosolvency, the free drug concentration available for complexation may be higher, leading to a synergistic improvement in solubility.
- Formation of soluble drug-ligand-cosolvent ternary complexes leads to synergistic improvement in solubility.
- Competition between the drug and cosolvent molecules for complexation with the ligand leads to a decrease in drug solubility.
- Decrease in apparent binding constant for drug-ligand in cosolvent can occur.

Thus, the impact of the combination of cosolvency and complexation on drug solubility needs to be evaluated on a case-by-case basis. Li et al.[68] provided a mathematical model for the total solubility by considering the free drug, drug-ligand complex, and drug-solvent complex. The authors assumed that the solubilizing power of the cosolvent remains unchanged even in the presence of a complexing agent; the decrease in binding constant can be empirically related to cosolvent concentration in an exponential fashion. They also assumed negligible complexation between cosolvent molecules and the complexing agent. Fluasterone solubility increased linearly with HP-β-CD in all ethanol-water mixtures. In the absence of HP-β-CD, an exponential increase in solubility with percentage w/w of ethanol in line with the log-linear model discussed previously was observed. In the presence of HP-β-CD, the drug solubility first decreased and then increased with an increase in percentage w/w of ethanol. Using the mathematical models developed for this system, the authors were able to explain the experimental data.

1.4.5.5 *Combined effect of complexation and micellar solubilization*

Several interactions (equilibria) can coexist in solubility experiments in aqueous solutions containing a

complexing agent, surfactant, and drug. They include competitive complexation of the drug and surfactant monomer with the complexing agent, equilibria between monomer and micelle, and solubilization of drug in the micelle. Yang et al.[73] presented experimental data and a semiquantitative analysis of solubilization of NSC-639829 in aqueous solutions of SLS and (SBE)7M-β-CD. Rao et al.[74] presented a mathematical model to study the combined effect of micellar solubilization and complexation. This model assumes that ligand molecules or drug-ligand complexes do not interact with micelles, and higher order drug-ligand or surfactant monomer:ligand complexes are not formed. The phase-solubility profile in the presence of both complexing agent and surfactant are dependent on a number of factors. However, it is possible for one to predict this behavior, provided that the binary interaction parameters among the drug, complexing agent, and surfactant are known.

The first step in understanding the various equilibria is to determine P_f, the surfactant concentration that is unbound to the complexing agent. This requires solving the following quadratic equation:

$$K_P \times [P_f]^2 + (1 + K_P \times L_T + K_D \times S_0 - K_P \times P_T) \times [P_f] - (P_T + P_T \times K_D \times S_0) = 0 \quad (1.74)$$

K_P and K_D are 1:1 binding constants for surfactant monomer-ligand and drug-ligand, respectively.

If P_f is found to be less than the known CMC value for the surfactant, the total drug solubility can be described as follows:

$$S_T = [D_f] + [D - CD] = S_o + \frac{K_D \times S_0 \times [L_T]}{1 + K_P \times [P_f] + K_D \times S_0} \quad (1.75)$$

According to this equation, the total solubility will be lower in the presence of surfactant, due to competition between the drug and surfactant monomer for the complexing agent (Scheme 1.2a).

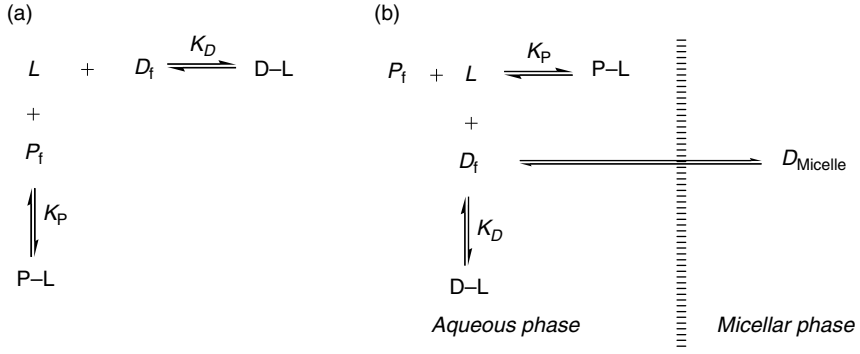
If P_f is greater than the known CMC value of the surfactant, the total solubility is defined as

$$S_T = S_o + K_M\left(P_T - \mathrm{CMC} - \frac{K_P \times \mathrm{CMC} \times [L_T]}{1 + K_P \times \mathrm{CMC} + K_D \times S_0}\right) + \frac{K_D \times S_0 \times [L_T]}{1 + K_P \times \mathrm{CMC} + K_D \times S_0} \quad (1.76)$$

According to this equation, the combined solubility is less than the sum of the solubility values in the presence of individual additives; that is, the complexing agent or surfactant (Scheme 1.2b). It is also possible that the combined solubility is less than individual solubility values. Yang et al.[73] showed that for NSC-639829, the combined solubility in aqueous solutions containing both SLS and (SBE)7M-β-CD is less than the sum of the solubility values in aqueous solutions containing SLS or (SBE)7M-β-CD.

1.5 EXPERIMENTAL DETERMINATION OF SOLUBILITY

The importance of knowing the solubility of a drug was discussed earlier in this chapter. As a consequence, the determination of solubility remains one of the most routinely conducted experiments during the drug development stage. Although these experiments are commonly perceived as simple, accurate solubility determination may be far more challenging. Pontolillo and Eganhouse[75] have noted a scatter of five log units in the experimentally determined solubility of dichlorodiphenyltrichloroethane (DDT). The AQUASOL database records a scatter of 10−30 times in the solubility of anthracene and fluoranthrene reported in the literature. These observations showcase the fact that getting an accurate determination of solubility is not a trivial task. Various aspects related to the solute and solvent of interest, experimental conditions, and the analytical



SCHEME 1.2 (a) Competitive complexation of drug and surfactant monomer to ligand ($P_f < \mathrm{CMC}$). (b) Competitive complexation and drug solubilization into micellar phase ($P_f > \mathrm{CMC}$). *Source: Redrawn based on Pontolillo J, Eganhouse E.* US geological survey, report 4201.

techniques employed, play important roles in the determination of actual solubility. This section covers the various methods that have been conventionally employed for solubility determination. In addition, several factors that may affect the solubility and its measurement will be discussed.

All methods used for solubility determination primarily consist of two parts: saturating the solvent with the solute and measuring the amount of solute present per unit of the solvent at that saturation state. In simple terms, the first part involves providing ample exposure of the solute to the solvent such that the dissolved solute molecules are in dynamic equilibrium with the undissolved molecules. A variety of analytical methods can then be used to measure the concentration of dissolved solute.

Before designing any solubility experiments, it is essential to identify whether the requirement is to determine thermodynamic equilibrium solubility or a kinetic dissolution rate. These concepts were discussed in the previous section. The need to determine one over the other depends upon the application of the results obtained. If the goal is to prepare a liquid formulation, the equilibrium solubility information is more meaningful. However, dissolution rates are needed to access the release of drug from the formulation and its subsequent absorption in the gastrointestinal tract.

The conventional approaches for measuring equilibrium solubility are based on the phase-solubility technique proposed by Higuchi and Connors.[54] An excess amount of the solute is placed with the solvent in a closed container and shaken for a sufficient length of time. Immersing sealed containers in a constant temperature bath may be used to control the temperature of the system. Once equilibrated, the sample is phase-separated to remove undissolved solute, and then analyzed quantitatively for dissolved solute content using nephlometry, UV absorbance, or potentiometric measurements.[76,77] It is necessary to maintain the same temperature used for equilibration during the phase-separation process. The main disadvantage of this method is that it can be time consuming, especially when there is a desire to run solubility experiments on a large number of compounds during the early discovery phase. Therefore, many variants of kinetic solubility experiments have been developed to rapidly screen a large number of compounds by either dissolving them in a solvent such as dimethyl sulfoxide (DMSO) or using an unidentified crystalline or amorphous phase. Various considerations are required while determining the solubility using this approach. Although the specific requirements can (and will) vary case by case, some general points are discussed next.

1.5.1 Stability of solute and solvent

It is essential that both the solute and solvent remain chemically and physically stable during the equilibration. The solute undergoing chemical degradation can result in more or less soluble products and provide misleading results. If the solvent is aqueous, hydrolysis of the solute may occur, generally resulting in the formation of more soluble degradation products. If the analytical method is not highly specific, such as UV-spectrophotometry, the measured solubility will be artificially higher than the actual solubility. The physical stability of both the solute and solvent must also be adequately considered. A hygroscopic solvent, like polyols, may pick up moisture during equilibration, which will result in a compositional change. For such solvents, care must be taken to protect against exposure to moisture. Similarly, while dealing with volatile solvents like low-molecular-weight alcohols, the container must be tightly closed to avoid solvent loss.

The physical form of the solute is an important determinant of its solubility. Polymorphic transitions, including conversion to or from hydrates/solvates, may occur during the equilibration process. Thus, it is essential to determine the form of the undissolved solute after equilibrium is achieved using a solid-state characterization technique like powder X-ray diffractometry. Determination of equilibrium solubility of a metastable form is difficult due to the conversion to the more stable form,[78] and this is particularly the case for compounds that are anhydrates or lower-order hydrates that convert to the hydrate formed upon equilibration. The best that one can do is to determine apparent solubility, often assumed to be the highest concentration observed during a solubility experiment.[79] Alternatively, the ratio of intrinsic dissolution rates for the two forms at identical hydrodynamic conditions, along with the solubility of the stable form, can be used to determine the solubility of the metastable form.[80]

On similar grounds, for ionizable solutes, it must be determined whether the solute exists in the free form or as a salt. If the solubility is to be determined at a specific pH, adequate measures must be taken to maintain the pH in the desired range using buffers. The choice of buffer should be based on the pH range desired and the selected buffer should have adequate buffer strength in that range. Monitoring and adjusting the pH of the system are essential, and the frequency of doing so strictly depends on the strength of the buffer used, ionic strength of the solute, and its dissolution rate. In addition to the pH, the ionic strength of the media affects the equilibrium solubility and must be maintained during the equilibration using common salts such as sodium or potassium chloride. The salt

selected must not contain the same ion present in the solute to avoid the common-ion effect.[81,82] Buffer components and salts used to adjust ionic strength may themselves form salts with ionizable drugs that have limited solubility. Thus, the identity of all insoluble residues must be confirmed, and the absence of buffer components or counterions of salts used to adjust ionic strength must be demonstrated.

1.5.2 Shakers and containers

Proper wetting of the solute has to be ensured by choosing an appropriate shaker. The use of several types of shakers has been reported in the literature, including end-over-end shakers, wrist action shakers, magnetic stirrers, vortexers, etc. The selection of a shaker depends on the volume and viscosity of the solvent used. Generally, end-over-end shakers provide good solute-solvent exposure for small volumes (>20 mL) of solvents that have viscosity similar to that of water. Industrial-size shakers are available for the manufacture of solution formulations. The compatibility of the solute and the solvent with the container is essential. Glass is generally the material of choice, at least for small-scale experiments, as it is relatively chemically inert. Protection from light exposure for photo-labile systems can be achieved by either using amber vials or covering the container with an opaque material such as aluminum foil.

1.5.3 Presence of excess undissolved solute

The presence of some undissolved solute during the entire length of equilibration is essential. However, a lot of undissolved solute must be avoided, as there have been several reports claiming the dependence of the equilibrium solubility on the amount of excess solute present during equilibration. Wang et al.[82] noticed that the solubility of a diprotic weak base depended on the amount of hydrochloride salt added. More recently, Kawakami et al.[83] reported that when 40 μg of indomethacin was added to 1 mL pH 5 or 6 citrate buffer, a much higher solubility was obtained compared to 5 mg added to the same solution, while a reverse trend was noticed in pH 6.5 and 7.0 phosphate buffer.

Supersaturation is another cause of imprecise solubility data. Ledwidge and Corrigan[84] reported that the self-association near pH_{max} reduces the rate of nucleation of the final salt, and that when the free drug is used as the starting material, higher levels of supersaturation are achieved.

Caution must be taken if the solute contains soluble impurities. These impurities can influence the property of the media and, consequently, the solubility of the solute. They may also interfere with the technique used for solubility analysis. This is of particular concern if the analytical technique used is nonselective. The presence of impurities has been shown to affect the solubility of 7-(2-hydroxypropyltheophylline).[85]

1.5.4 Determination of equilibrium

Information from prior experiments can be useful for gauging the time to achieve equilibrium. If no prior information is available, several samplings may be required to establish the time it takes to reach equilibrium. The periodicity of withdrawing sample depends on the dissolution rate of the solute. In order to minimize sampling, the interval can progressively be doubled. In other words, if the first sampling is performed 2 h following the start of the experiment, the second sampling can be done after 4 h, the third after 8 h, and so on. The sampling should be continued until the solubility for the last two sampling points is equal, indicating the attainment of equilibrium. The last sampling point can be conservatively treated as the time required for the equilibration. While conducting this exercise, it is important not to replace the solvent removed during sampling, as it will push the system out of equilibrium.

1.5.5 Phase separation

Once the equilibrium between the dissolved and undissolved solute is achieved, the dissolved phase has to be separated and analyzed for solute concentration. Several methods can be used to phase-separate the system. Filtration provides a convenient way to separate the dissolved phase and the undissolved solute. Appropriate filters can be chosen from a variety of configurations available to suit the need of the experiment. Commercial filters are available in a variety of sizes, pore dimensions, and membrane materials. The compatibility of the filter with both the solute and solvent must be determined. Strongly organic solvent systems can leach out components from the filter membrane or its housing. On the other hand, some solutes can adsorb on the filter assembly, which can result in solubility artifacts. An example is amiodarone, which adsorbs onto some polyvinyl-based filters. As a general practice, the filter system must be "rinsed" by passing through the solution two to three times before collection. This ensures that the system is saturated with the solute and solvent prior to sampling and will reduce artifacts.

Phase separation can also be achieved by centrifugation. This technique is particularly useful if both the

solute and the solvent are liquids. Following phase separation, the clear supernatant is collected, leaving behind the undissolved solute. Considerations must be made for possible compatibility issues of the solute or solvent with centrifuge tubes. Glass tubes are preferred whenever possible to minimize adsorption issues. Other factors include the speed and length of centrifugation, which depend on the properties of the solute and solvent.

Other methods for phase separation include the evaporative collection of solvent and selective adsorption of dissolved phase. These methods, although they can be useful in some situations, are less practical to carry out. Whichever method of phase separation is used, control over temperature during the separation process is essential. Prolonged contact of undissolved solids with the supernatant-saturated solution during the separation process at a temperature other than that at which equilibration was performed will likely result in underestimation or overestimation of solubility, depending on the temperature difference between the equilibration process and the phase-separation process.

1.5.6 Determination of solute content in the dissolved phase

Once the dissolved phase is separated, the solute concentration is determined analytically. Several techniques can be used, including spectrophotometry, gravimetry, and pH measurement. Spectrophotometry is the most commonly employed technique based on its simplicity, accuracy, and chemical specificity. This technique is based on the well-known Beer's law, according to which the absorbance of a chemical species is directly proportional to its concentration. A simple analysis can be performed by plotting a standard curve using known concentrations of the solute, followed by quantification of the unknown solute concentration in the solution. The solution may be diluted if needed, to fit into the concentration range of the standard curve.

In order to use this technique, it is essential that the analyte molecules contain a chromophore; that is, it is spectrally active in the UV or visible light region. If that is not the case, it can be chemically derivatized to make it spectrally active. The complexity of derivatization reactions depends upon the chemical structure of the solute, and in some cases, it may be a limiting factor.

It is a common practice to couple spectrophotometry with separation techniques, such as chromatography, to improve the specificity of the experiment. This approach is very useful in cases where either solvent or impurity/degradation products interfere with direct solute determination. It is particularly important at the early stages of development, where solute purity may not be known.

Gravimetric analysis involves the determination of the weight of dissolved solute in solution. This is primarily achieved either by solvent evaporation or solute precipitation using physical or chemical means. This approach is relatively simple, inexpensive, and does not require lengthy method-development steps. However, its broader application is marred by lack of sensitivity, selectivity, and practicality. Measurements of indirect properties, such as the pH of the system or colligative properties that can yield information on the concentration of solute in a given solution, also have been proposed.

1.5.7 Experimental conditions

The external conditions in which the experiment is carried out can have a huge impact on the solubility results. The effect of temperature on solubility has been detailed in the discussion following Eq. (1.11). Temperature may have a positive or a negative influence on the solubility, depending upon the solute and solvent properties. Pressure significantly affects the solubility of gases but generally has little influence on the solubility of liquids and solids. The presence of light may affect the chemical stability of photolabile solutes and may artificially affect their solubility.

Besides the thermodynamically correct phase-solubility method, the solubility can be ball-parked by using the synthetic method. According to this method, a known amount of solute is added to a known amount of solvent, followed by a mixing step. If the solute dissolves completely, more is added and the process continues until a stage is reached after which the added solute does not dissolve. Alternatively, if the solute does not dissolve completely after the first round, more solvent is added, and the process continues until the solute dissolves completely. The approximate solubility can be calculated by the knowledge of the amount of solvent required to completely dissolve a known quantity of the solute. This method provides a simple means for approximate solubility determination. It is particularly useful for solutes available in limited amounts, as the development of involved analytical method is not required. However, the approach lacks the sensitivity and selectivity of the phase-solubility approach.

References

1. Amidon GL, Lennernas H, Shah VP, Crison JR. *Pharm Res* 1995;**12**:413–20.
2. Gribbon P, Sewing A. *Drug Discovery Technologies* 2005;**10**:17–22.
3. Deshwal BR, Lee KC, Singh KC. *J Mol Liq* 2006;**123**:38–42.

4. Noyes AA, Whitney WR. *J Am Chem Soc* 1897;**19**:930–4.
5. Hamlin WE, Northam JI, Wagner JG. *J Pharm Sci* 1965;**54**:1651–3.
6. Sanghvi R. *Drug solublization using N-methyl pyrrolidone: efficiency and mechanism, PhD thesis*. The University of Arizona; 2006.
7. Jain P, Yalkowsky SH. *Int J Pharm* 2007;**342**:1–5.
8. Hildebrand JH, Prausnitz JH, Scott RL. *Regular and related solutions*. New York: Van Nostrand Reinhold; 1970.
9. Prausnitz JM, Lichtenhaler RL, Gomez de Azevado E. *Molecular thermodynamics of fluid phase equilibria*. 2nd ed. Englewood Cliffs (NJ): Prentice Hall; 1986.
10. Grant DJW, Mendizadah M, Chow AHL, Fairbrother JE. *Int J Pharm* 1984;**18**:25–38.
11. Mishra DS. *Solubility of organic compounds in nonaqueous systems, PhD thesis*. The University of Arizona; 1988.
12. Mishra DS, Yalkowsky SH. *Ind Eng Chem Res* 1990;**29**:2278–83.
13. Walden P. *Zeitschrift fur Angewandte Physik und Chemie* 1908;**14**:713–28.
14. Martin E, Yalkowsky SH, Wells JE. *J Pharm Sci* 1979;**68**:565–8.
15. Dannenfelser RM, Yalkowsky SH. *Ind Eng Chem Res* 1996;**35**:1483–6.
16. Jain A, Yang G, Yalkowsky SH. *Ind Eng Chem Res* 2004;**43**:4376–9.
17. Hildebrand JH, Scott RL. *The solubility of nonelectrolytes*. New York: Rheinhold Publishing Corporation; 1950.
18. Hansen CM, Beerbower A. *Encyclopedia of chemical technology*. 2nd ed. New York: Wiley; 1971.
19. Fedor R. *Polym Eng Sci* 1974;**14**(2):147–54.
20. Sanghvi R, Yalkowsky SH. *Ind Eng Chem Res* 2006;**45**:2856–61.
21. Martin A, Newburger J, Adeji AJ. *J Pharm Sci* 1979;**68**:iv.
22. Yalkowsky SH, Flynn GL, Amidon GL. *J Pharm Sci* 1972;**61**:983–4.
23. Yalkowsky SH, Amidon GL, Zografi G, Flynn GL. *J Pharm Sci* 1975;**64**:48–52.
24. Yalkowsky SH, Valvani SC, Amidon GL. *J Pharm Sci* 1976;**65**:1488–93.
25. Amidon GL, Yalkowsky SH, Leung S. *J Pharm Sci* 1974;**63**:1858–66.
26. Myrdal PB, Ward GH, Dannenfelser RM, Mishra DS, Yalkowsky SH. *Chemosphere* 1992;**24**:1047–61.
27. Myrdal PB, Ward GH, Simamora P, Yalkowsky SH. *SAR QSAR Environ Res* 1993;**1**:55–61.
28. Myrdal PB, Yalkowsky SH. *SAR QSAR Environ Res* 1994;**2**:17–28.
29. Myrdal PB, Manka A, Yalkowsky SH. *Chemosphere* 1995;**30**:1619–37.
30. Hansch C, Quinlan JE, Lawrence GL. *J Org Chem* 1968;**33**:347–50.
31. Jain N, Yalkowsky SH. *J Pharm Sci* 2001;**90**:234–52.
32. Sepassi K, Yalkowsky SH. *AAPS PharmSciTech* 2006;**7**(1). Article 26, http://www.aapspharmscitech.org.
33. Machatha SG, Yalkowsky SH. *Int J Pharm* 2005;**294**:185–92.
34. Katritzky AR, Jain R, Lomaka A, Petrukhin R, Maran U, Karelson M. *Cryst Growth Des* 2003;**1**:261–5.
35. Dearden JC. *Environ Toxicol Chem* 2003;**22**:1696–709.
36. Bergstrom CAS, Norinder U, Luthman K, Artursson P. *J Chem Inf Comput Sci* 2003;**43**:1177–85.
37. Jain A, Yalkowsky SH. *J Pharm Sci* 2006;**95**:2562–618.
38. Niimi A. *Water Res* 1991;**25**:1515–21.
39. Bruggemann R, Altschuh J. *QSAR Environ Toxicol* 1991;**4**:41–58.
40. Ran Y, Jain N, Yalkowsky SH. *J Chem Inf Comput Sci* 2001;**41**:354–7.
41. Ran Y, He Y, Yang G, Johnson JLH, Yalkowsky SH. *Chemosphere* 2002;**48**:487–509.
42. Jain N, Yang G, Machatha SG, Yalkowsky SH. *Int J Pharm* 2006;**319**:169–71.
43. Dearden JC. *Expert Opin Drug Discov* 2006;**1**:31–52.
44. Jorgensen WL, Duffy EM. *Adv Drug Deliv Rev* 2002;**54**:355–66.
45. Abraham MH, Li J. *J Pharm Sci* 1999;**88**:868–80.
46. Taskinen J, Yliruusi J. *Adv Drug Deliv Rev* 2003;**55**:1163–83.
47. Gould PL. *Int J Pharm* 1986;**33**:201–17.
48. Serajuddin ATM, Pudipeddi M. *Handbook of pharmaceutical salts*. Wiley-VCH; 2002. p. 135–60.
49. Streng WH, His SK, Helms PE, Tan HGH. *J Pharm Sci* 1984;**73**:1679–84.
50. Chowhan ZT. *J Pharm Sci* 1978;**67**:1257–60.
51. Bogardus JB, Blackwood RK. *J Pharm Sci* 1979;**68**:188–94.
52. Maurin MB, Grant DJW, Stahl PH. *Handbook of pharmaceutical salts*. Wiley-VCH; 2002. p. 9–18.
53. Repta AJ. *Techniques of solubilization of drugs*. New York: Marcel Dekker Inc; 1981. p. 135–58.
54. Higuchi T, Connors KA. *Adv Anal Chem* 1965;117.
55. Higuchi T, Lach J. *J Am Pharm Assoc* 1954;**43**:349.
56. Higuchi T, Pitman IH. *J Pharm Sci* 1973;**62**:55.
57. Desiraju GR. *CrystEngComm* 2003;**5**:466–7.
58. Dunitz JD. *CrystEngComm* 2003;**5**:506.
59. Vishweshwar P, McMahon JA, Bis JA, Zaworotko MJ. *J Pharm Sci* 2006;**95**:499–516.
60. Nehl SJ, Rodriguez-Spong B, Rodriguez-Hornedo N. *Cryst Growth Des* 2006;**592**:600.
61. Yalkowsky SH. *Solubility and solubilization in aqueous media*. Oxford (UK): Oxford University Press; 1999.
62. Gould PL, Goodman M, Hanson PA. *Int J Pharm* 1984;**19**:149–59.
63. Yalkowsky SH, Amidon GL, Zografi G, Flynn GL. *J Pharm Sci* 1972;**61**:48–52.
64. Yalkowsky SH, Roseman TJ. *Techniques of solubilization of drugs*. New York: Marcel Dekker Inc; 1981. p. 91–134.
65. Yalkowsky SH, Robino JT. *J Pharm Sci* 1983;**72**:1014–17.
66. Millard JW, Alvarez-Nunez FA, Yalkowsky SH. *Int J Pharm* 2002;**245**:153–66.
67. Kramer SF, Flynn GL. *J Pharm Sci* 1972;**61**:1896–904.
68. Li P, Tabibi E, Yalkowsky SH. *J Pharm Sci* 1999;**88**:945–7.
69. Tinwalla AY, Hoesterey BL, Xiang T, Lim K, Anderson BD. *Pharm Res* 1993;**10**:1136–43.
70. Johnson MD, Hoesterey BL, Anderson BD. *J Pharm Sci* 1994;**83**:1142–6.
71. Okimoto K, Rajeswki RA, Uekama K, Jona JA, Stella VJ. *Pharm Res* 1996;**13**:256–64.
72. Rao VM, Stella VJ. *J Pharm Sci* 2003;**92**:927–32.
73. Yang G, Jain N, Yalkowsky SH. *Int J Pharm* 2004;**269**:141–8.
74. Rao VM, Nerurkar M, Pinnamaneni S, Raghavan K. *Int J Pharm* 2006;**319**:98–106.
75. Pontolillo J, Eganhouse RP. *US Geological Survey, water resources investigations report 01-4201, 2001*.
76. Glomme A, Marz J, Dressman JB. *J Pharm Sci* 2005;**94**:1–16.
77. Fligge TA, Shuler A. *J Pharm Biomed Anal* 2006;**42**:449–54.
78. Khankari RK, Grant DJW. *Thermochimica Acta* 1995;**248**:61–79.
79. Morris KR, Rodriguez-Hornedo N. *Encyclopedia of pharmaceutical technology* 1993;**393**.
80. Anderson BD, Flora KP. *The practice of medicinal chemistry* 1996;739–54.
81. Pudipeddi M, Serajuddin ATM, Grant DJW, Stahl PH. *Handbook of pharmaceutical salts*. Wiley-VCH; 2002. p. 19–39.
82. Wang Z, Burrell LS, Lambert WJ. *J Pharm Sci* 2002;**91**:1445–55.
83. Kawakami K, Miyoshi K, Ida Y. *Pharmacol Res* 2005;**22**:1537–43.
84. Ledwidge MT, Corrigan OI. *Int J Pharm* 1998;**174**:187–200.
85. Giron D, Goldbronn C. *Analusis* 1979;**7**:109–26.

CHAPTER

2

Crystalline and Amorphous Solids

G.G.Z. Zhang and D. Zhou

AbbVie Inc., North Chicago, IL, United States

2.1 INTRODUCTION

Of the several states of matter in which a substance can reside, the solid state is most commonly encountered; therefore, it is the most important and relevant state for pharmaceutical development. Most of the pharmaceutical products on the market or formulations presently being developed are in the solid dosage form. Even when a product is marketed or developed as a solution or a semisolid formulation, a solid is usually selected and manufactured as an active pharmaceutical ingredient (API) for such reasons as the ability to crystallize (and therefore to be purified), ease of handling, and better chemical stability in comparison with liquids. Therefore, an understanding of the various solid forms that may occur, as well as the rational selection of solid forms for development, are critical to the facile development of a particular chemical entity. Many books[1–7] and special journal issues[8–12] are devoted to this topic. This chapter, due to space limitations, provides only a brief introduction to the fundamental principles and practical aspects of pharmaceutical solids. Interested readers are encouraged to read the books and reviews cited here.

2.2 DEFINITIONS AND CATEGORIZATION OF SOLIDS

The various types of pharmaceutical solids are shown in Fig. 2.1. Solids can show differences either externally or internally. External differences are considered as the shape, habit, or morphology of the particles, where the internal structures that make up the solid particles remain the same. Although important for pharmaceutical development, these elements are not discussed in this chapter. The focus of this chapter is on solids with distinct differences in their internal structures.

Based on the degree of long-range order/periodicity, solids are categorized into three groups. Amorphous phases are those solids that do not exhibit long-range order in any of the three physical dimensions. However, short-range order could exist for amorphous solids. Because of the importance of this class of solids to pharmaceutical development, it is discussed in detail in Section 2.7 of this chapter. If materials have long-range order in only one or two dimensions, they are liquid crystalline in nature. Liquid crystalline materials can be further categorized based on the number of components contained therein, as is the case for crystalline solids. Since liquid crystals, with properties intermediate to conventional liquids and three-dimensional solids, are not frequently encountered, they will not be discussed in detail. The vast majority of pharmaceutical solids fall into the category of crystalline solids because they exhibit long-range order in all three dimensions.

Crystalline solids can be further categorized into various subtypes based on the number of components that make up the solid internally, in a homogeneous fashion. The solid could be composed of the drug alone, or as adducts with one (binary), two (ternary), three (quaternary), other chemical species. Although the number of other chemical species, apart from the drug itself, can increase without limit, it usually is a relatively low integer.

When the overall chemical composition of solids is the same, they can be different in internal structures. The ability of a substance to exist as two or more crystalline phases that have different arrangements or conformations of the molecules in a crystalline lattice is called *polymorphism*, and these different solids are termed *polymorphs*. One point to emphasize is that, according to

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00002-9

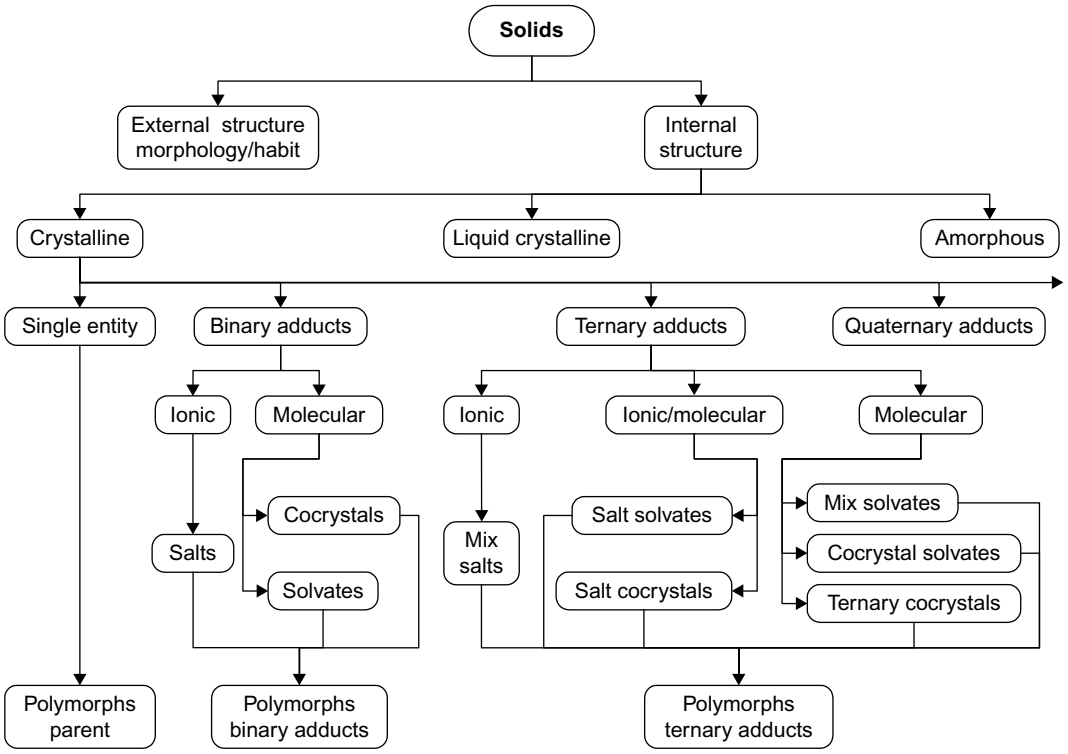


FIGURE 2.1 Categorization of pharmaceutical solids.

the strict definition of this term, different polymorphs are only different physically, not chemically. When these solids are melted or dissolved in solutions, they are exactly the same, both physically and chemically. However, in some special cases, the definition of polymorphism is extended slightly for convenience. There are many possible reasons for a substance to exhibit polymorphism, and sometimes it may be difficult to pinpoint a single one. These reasons include packing; conformation (which usually will exhibit different packing as well); hydrogen bonding pattern/motif; chirality (ie, racemic compound versus racemic conglomerate; the definition is extended here because the racemic conglomerate is an equal molar physical mixture of the enantiomerically pure crystals, and therefore is not a single phase); and tautomerism (again, the definition is extended because different tautomers are chemically not equivalent). Polymorphism is relevant not only to solids composed of single components, but also to adducts of multiple components (binary, ternary, quaternary, etc.). However, different polymorphs should have the same chemical composition, qualitatively and quantitatively, apart from the extensions of the definition mentioned here.

Adducts are formed where multiple chemical species are present in the crystalline lattices, again in a homogeneous fashion. Adducts can be categorized further based on the ionization states of these species: ionic, molecular, or ionic/molecular. The same subcategorization described in this chapter applies to quaternary and higher adducts.

An ionic adduct is one that is made of ions only—cations and anions. For binary adducts, these are called *salts.* Various acids and bases can pair up to form different salts. The same acid/base pair can form salts of different stoichiometries. For ternary adducts, these are called *mixed salts,* where one additional acid or base is involved in salt formation.

A molecular adduct is one that is made of neutral molecules. Depending on the physical states of the additional components, it can be a solvate or a cocrystal. When this additional component is a liquid in ambient conditions in its pure form, the molecular adduct is called a *solvate.* If the solvent is water, it is called a *hydrate.* However, when the additional component is a solid in ambient conditions, the molecular adduct is called a *cocrystal.* The boundary separating these two types of solids is somewhat arbitrary. Therefore, many published articles include solvates as cocrystals. However, there are distinct differences between the two types of solid adducts. In some situations, it may be helpful to distinguish them.

An ionic/molecular adduct is one that is made of both ionic and neutral molecules. This type of solid only exists for adducts with three or more

components. One important thing to clarify here is that an acid or a base is not considered the same component as its ion. Thus, a ternary ionic/molecular adduct is composed of one salt molecule (two components) and one neutral molecule. Again, based on the physical state of the neutral component, it could be a salt solvate or a salt cocrystal. Unlike the previous categorization of solids,[13] this categorization system does not consider the structural details within the solids; therefore, it is widely applicable and easier to use.

Cocrystals, although they have physically existed for some time and were formerly termed *molecular complexes*, are a relatively new concept. The definition of the term has been a topic of discussion in recent years.[14–19] Based on these discussions, and the authors' personal viewpoints, the term *cocrystal* is here defined as "structurally homogeneous crystalline molecular adducts, made from components that are apparently neutral, that are by themselves solids at ambient conditions. The components are held together by interactions other than covalent or ionic bonds (hydrogen bonding, $\pi-\pi$, van der Waals, charge-transfer, halogen–halogen, etc.)."

Most adducts are stoichiometric in nature; that is, there is a defined ratio between the components. However, some are nonstoichiometric,[20–25] such as cromolyn sodium and cefazolin sodium hydrates.[22,23]

In view of the importance of salts to pharmaceutical development, they will be discussed separately in chapter "API Solid-Form Screening and Selection." This chapter will include discussions on the other crystalline solids (ie, polymorphs, solvates/hydrates, cocrystals, and amorphous solids).

2.3 THERMODYNAMICS AND PHASE DIAGRAMS

Thermodynamics is one of the most important aspects of understanding pharmaceutical solids. Phase diagrams are usually constructed to express the thermodynamic relationships among various solid phases, qualitatively or quantitatively. Depending on the types of solids, phase diagrams are constructed with respect to variables of different physical significance.

2.3.1 Polymorphs

2.3.1.1 Enantiotropy and monotropy

When examining a pair of polymorphs, Polymorph I and Polymorph II, the stability relationship between the two polymorphs is determined *exclusively* by their free energy differences at different temperatures. The free energy of a particular solid is expressed in Eq. (2.1):

$$G = H - TS \tag{2.1}$$

where:

G is the Gibbs free energy,
H is the enthalpy,
T is the temperature,
S is the entropy.

Thus, the free energy for the transition from Polymorph I to Polymorph II is expressed as

$$\begin{aligned}\Delta G_{\mathrm{I}\to\mathrm{II}} &= G_{\mathrm{II}} - G_{\mathrm{I}} = (H_{\mathrm{II}} - TS_{\mathrm{II}}) - (H_{\mathrm{I}} - TS_{\mathrm{I}}) \\ &= \Delta H_{\mathrm{I}\to\mathrm{II}} - T\Delta S_{\mathrm{I}\to\mathrm{II}}\end{aligned} \tag{2.2}$$

$$\Delta H_{\mathrm{I}\to\mathrm{II}} = H_{\mathrm{II}} - H_{\mathrm{I}} \quad \text{and} \quad \Delta S_{\mathrm{I}\to\mathrm{II}} = S_{\mathrm{II}} - S_{\mathrm{I}}$$

At any particular temperature, three different situations exist:

1. $\Delta G_{\mathrm{I}\to\mathrm{II}} < 0$: Polymorph II has lower free energy and is therefore more stable than Polymorph I. The transition from Polymorph I to Polymorph II is a spontaneous process.
2. $\Delta G_{\mathrm{I}\to\mathrm{II}} > 0$: Polymorph II has higher free energy and is therefore less stable than Polymorph I. The transition from Polymorph I to Polymorph II is not a spontaneous process, but the transition from Polymorph II to Polymorph I is.
3. $\Delta G_{\mathrm{I}\to\mathrm{II}} = 0$: Polymorph I and Polymorph II have the same free energy. Therefore, both polymorphs have equal stability. There will be no transition between the two polymorphs.

As shown in Eq. (2.2), the free energy difference between the two polymorphs changes with temperature. The temperature at which the two polymorphs have equal stability is defined as the transition temperature (T_t).

If T_t is located below the melting points of both polymorphs, the two polymorphs are said to be *enantiotropes*, and the polymorphic system is said to exhibit enantiotropy or to be enantiotropic in nature. A representative phase diagram for an enantiotropic polymorphic system is shown in Fig. 2.2a. Below T_t, Polymorph I is more stable. Above T_t, Polymorph II is more stable. In this type of system, the melting point of Polymorph I is lower than that of Polymorph II, although the melting point of Polymorph I may not be experimentally accessible because of facile solid-solid transitions at lower temperatures.

If T_t is located above the melting points of both polymorphs, the two polymorphs are said to be *monotropes*, and the polymorphic system is said to exhibit monotropy or to be monotropic in nature. A representative

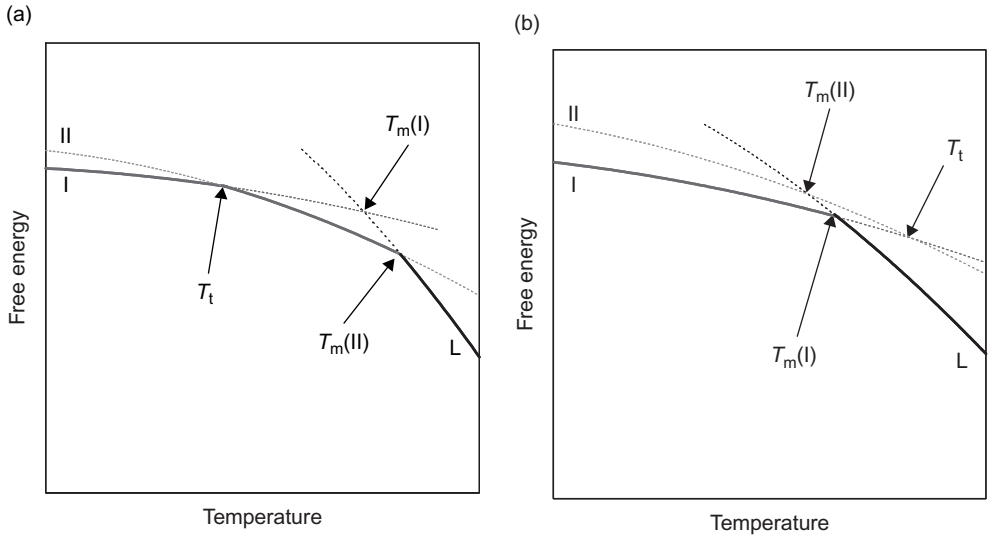


FIGURE 2.2 Thermodynamic phase diagrams of polymorphs. (a) Enantiotropy and (b) Monotropy.

phase diagram for a monotropic polymorphic system is shown in Fig. 2.2b. Throughout the temperature range, Polymorph I is more stable. Here, T_t is a hypothetical temperature, which is experimentally not accessible. In this type of system, the melting point of Polymorph I is higher than that of Polymorph II, although that of Polymorph II may not be experimentally accessible because of facile solid-solid transitions at lower temperatures.

In constructing phase diagrams, the convention is to plot data for the stable phase using a solid line and to use dotted lines to represent metastable phases, as shown in Fig. 2.2.

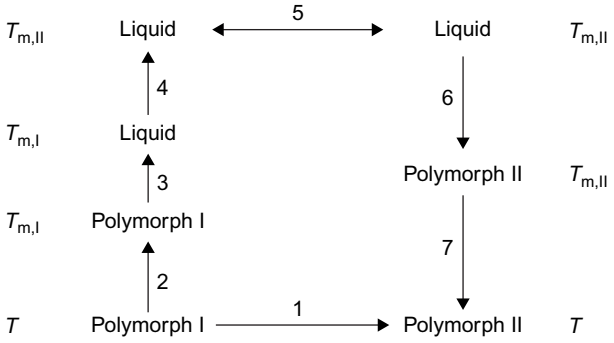


FIGURE 2.3 Hess cycle for polymorphic transition.

2.3.1.2 *Methods of determining stability relationships between polymorphs*

Many methods are available for determining stability relationships between polymorphs. Qualitative methods only determine the relationship with no precise knowledge of T_t. Quantitative methods, on the other hand, determine T_t first. The T_t is then compared to the melting points of each polymorphic form, and the stability relationship is thus defined. The relevant temperatures in which the stability relationship is of interest usually ranges from ambient temperatures to the melting point. However, in certain cases, the pertinent lower end of this range can extend below ambient temperatures—for instance, down to −50°C, which is relevant to freeze-drying processes.

2.3.1.2.1 Quantitative methods

2.3.1.2.1.1 Using heat of fusion data[26] When heats of fusion can be experimentally determined, the stability relationship of the two polymorphs can be determined by calculating T_t and comparing the result to the melting points. If T_t is less than the lower melting point of the two polymorphs, the system is enantiotropic. If the calculated T_t is more than the higher melting point of the two polymorphs, the system is monotropic.

Calculation of T_t begins with the design of a Hess cycle for the polymorphic transition, as shown in Fig. 2.3, assuming that Polymorph II has a higher melting point. Process 1 represents the polymorphic

transition I→II at the temperature of interest. Alternatively, the same transition can be achieved by the series of Processes 2−7. The thermodynamic quantities, ΔH and ΔS, for each process are listed in Table 2.1.

Since Process 1 is equivalent to the sum of Processes 2−7, the enthalpy and entropy of Process 1 are the sums of the enthalpies and entropies of Processes 2−7 (Eq. 2.3):

$$\begin{aligned}
\Delta H_{\mathrm{I}\to\mathrm{II}} &= \int_{T}^{T_{\mathrm{m,I}}} C_{\mathrm{p,I}}dT + \Delta H_{\mathrm{m,I}} + \int_{T_{\mathrm{m,I}}}^{T_{\mathrm{m,II}}} C_{\mathrm{p,L}}dT + 0 \\
&\quad - \Delta H_{\mathrm{m,II}} - \int_{T}^{T_{\mathrm{m,II}}} C_{\mathrm{p,II}}dT \\
&= (\Delta H_{\mathrm{m,I}} - \Delta H_{\mathrm{m,II}}) - \int_{T}^{T_{\mathrm{m,I}}} \Delta C_{\mathrm{p}}dT \\
&\quad - \int_{T_{\mathrm{m,I}}}^{T_{\mathrm{m,II}}} (C_{\mathrm{p,II}} - C_{\mathrm{p,L}})dT \\
&\approx (\Delta H_{\mathrm{m,I}} - \Delta H_{\mathrm{m,II}}) - \int_{T}^{T_{\mathrm{m,I}}} \Delta C_{\mathrm{p}}dT \\
&\approx \Delta H_{\mathrm{m,I}} - \Delta H_{\mathrm{m,II}} \\
\Delta S_{\mathrm{I}\to\mathrm{II}} &= \int_{T}^{T_{\mathrm{m,I}}} \frac{C_{\mathrm{p,I}}}{T}dT + \Delta S_{\mathrm{m,I}} + \int_{T_{\mathrm{m,I}}}^{T_{\mathrm{m,II}}} \frac{C_{\mathrm{p,L}}}{T}dT + 0 \\
&\quad - \Delta S_{\mathrm{m,II}} - \int_{T}^{T_{\mathrm{m,II}}} \frac{C_{\mathrm{p,II}}}{T}dT \\
&= \left(\frac{\Delta H_{\mathrm{m,I}}}{T_{\mathrm{m,I}}} - \frac{\Delta H_{\mathrm{m,II}}}{T_{\mathrm{m,II}}}\right) - \int_{T}^{T_{\mathrm{m,I}}} \frac{\Delta C_{\mathrm{p}}}{T}dT \\
&\quad - \int_{T_{\mathrm{m,I}}}^{T_{\mathrm{m,II}}} \left(\frac{C_{\mathrm{p,II}}}{T} - \frac{C_{\mathrm{p,L}}}{T}\right)dT \\
&\approx \left(\frac{\Delta H_{\mathrm{m,I}}}{T_{\mathrm{m,I}}} - \frac{\Delta H_{\mathrm{m,II}}}{T_{\mathrm{m,II}}}\right) - \int_{T}^{T_{\mathrm{m,I}}} \frac{\Delta C_{\mathrm{p}}}{T}dT \\
&\approx \frac{\Delta H_{\mathrm{m,I}}}{T_{\mathrm{m,I}}} - \frac{\Delta H_{\mathrm{m,II}}}{T_{\mathrm{m,II}}}
\end{aligned} \tag{2.3}$$

The free energy of Process 1 can then be calculated from its enthalpy and entropy (Eq. 2.4):

$$\begin{aligned}
\Delta G_{\mathrm{I}\to\mathrm{II}} &= \Delta H_{\mathrm{I}\to\mathrm{II}} - T\Delta S_{\mathrm{I}\to\mathrm{II}} \\
&= (\Delta H_{\mathrm{m,I}} - \Delta H_{\mathrm{m,II}}) - \int_{T}^{T_{\mathrm{m,I}}} \Delta C_{\mathrm{p}}dT \\
&\quad - \int_{T_{\mathrm{m,I}}}^{T_{\mathrm{m,II}}} (C_{\mathrm{p,II}} - C_{\mathrm{p,L}})dT \\
&\quad - T\left(\frac{\Delta H_{\mathrm{m,I}}}{T_{\mathrm{m,I}}} - \frac{\Delta H_{\mathrm{m,II}}}{T_{\mathrm{m,II}}}\right) + T\int_{T}^{T_{\mathrm{m,I}}} \frac{\Delta C_{\mathrm{p}}}{T}dT \\
&\quad + T\int_{T_{\mathrm{m,I}}}^{T_{\mathrm{m,II}}} \left(\frac{C_{\mathrm{p,II}}}{T} - \frac{C_{\mathrm{p,L}}}{T}\right)dT \\
&= \left[\Delta H_{\mathrm{m,I}}\left(1 - \frac{T}{T_{\mathrm{m,I}}}\right) - \Delta H_{\mathrm{m,II}}\left(1 - \frac{T}{T_{\mathrm{m,I}}}\right)\right] \\
&\quad - \left(\int_{T}^{T_{\mathrm{m,I}}} \Delta C_{\mathrm{p}}dT - T\int_{T}^{T_{\mathrm{m,I}}} \frac{\Delta C_{\mathrm{p}}}{T}dT\right) \\
&\quad - \left[\int_{T_{\mathrm{m,I}}}^{T_{\mathrm{m,II}}} (C_{\mathrm{p,II}} - C_{\mathrm{p,L}})dT - T\int_{T_{\mathrm{m,I}}}^{T_{\mathrm{m,II}}} \left(\frac{C_{\mathrm{p,II}}}{T} - \frac{C_{\mathrm{p,L}}}{T}\right)dT\right] \\
&\approx \left[\Delta H_{\mathrm{m,I}}\left(1 - \frac{T}{T_{\mathrm{m,I}}}\right) - \Delta H_{\mathrm{m,II}}\left(1 - \frac{T}{T_{\mathrm{m,II}}}\right)\right] \\
&\quad - \left(\int_{T}^{T_{\mathrm{m,I}}} \Delta C_{\mathrm{p}}dT - T\int_{T}^{T_{\mathrm{m,I}}} \frac{\Delta C_{\mathrm{p}}}{T}dT\right) \\
&\approx \Delta H_{\mathrm{m,I}}\left(1 - \frac{T}{T_{\mathrm{m,I}}}\right) - \Delta H_{\mathrm{m,II}}\left(1 - \frac{T}{T_{\mathrm{m,II}}}\right)
\end{aligned} \tag{2.4}$$

In Eqs. (2.2)−(2.4), the first approximation results from ignoring the heat capacity difference between Polymorph II and liquid, while the second approximation results from ignoring the heat capacity difference between Polymorph I and II.

Set $\Delta G_{\mathrm{I}\to\mathrm{II}}$ to zero, and the solution is the transition temperature. The caution to be exercised in this approach is that the solids of both polymorphs need to be highly crystalline; consequently, the melting points and heats of fusion can be accurately determined.

TABLE 2.1 Thermodynamic Quantities of Individual Processes in the Hess Cycle of a Polymorphic Transition

Process	Nature of the process	ΔH	ΔS
1	Polymorph transition from I to II	$\Delta H_{\mathrm{I}\rightarrow\mathrm{II}}$	$\Delta S_{\mathrm{I}\rightarrow\mathrm{II}}$
2	Raise temperature of Polymorph I from T to its melting point	$\int_{T}^{T_{\mathrm{m,I}}} C_{\mathrm{p,I}}\,dT$	$\int_{T}^{T_{\mathrm{m,I}}} \frac{C_{\mathrm{p,I}}}{T}\,dT$
3	Melt Polymorph I to liquid at its melting point	$\Delta H_{\mathrm{m,I}}$	$\Delta S_{\mathrm{m,I}} = \frac{\Delta H_{\mathrm{m,I}}}{T_{\mathrm{m,I}}}$
4	Raise the temperature of the liquid from melting point of Polymorph I to that of Polymorph II	$\int_{T_{\mathrm{m,I}}}^{T_{\mathrm{m,II}}} C_{\mathrm{p,L}}\,dT$	$\int_{T_{\mathrm{m,I}}}^{T_{\mathrm{m,II}}} \frac{C_{\mathrm{p,L}}}{T}\,dT$
5	Maintain the liquid at the melting point of Polymorph II	0	0
6	Crystallize Polymorph II from liquid at its melting point	$-\Delta H_{\mathrm{m,II}}$	$-\Delta S_{\mathrm{m,II}} = -\frac{\Delta H_{\mathrm{m,II}}}{T_{\mathrm{m,II}}}$
7	Lower temperature of Polymorph II to T from its melting point	$-\int_{T}^{T_{\mathrm{m,II}}} C_{\mathrm{p,II}}\,dT$	$-\int_{T}^{T_{\mathrm{m,II}}} \frac{C_{\mathrm{p,II}}}{T}\,dT$

2.3.1.2.1.2 Using eutectic fusion data Calculation of T_t using heat of fusion data requires experimentally accessible data. Very often, compounds chemically decompose upon melting or physically transform before melting. In these cases, reliable heat of fusion data is not available. However, eutectic melting can be employed and eutectic melting data can be used to calculate the free energy difference between the polymorphs.[27] This method has been successfully applied to complicated polymorphic systems such as ROY,[28,29] chiral drugs such as tazofelone,[27,30] amino acids such as glycine,[27] and sugars such as D-mannitol.[27] Due to limitations of space and the complexity of derivation, interested readers are encouraged to study the original articles referenced here for derivations and applications.

2.3.1.2.1.3 Using solubility/intrinsic dissolution rate data The solubility of a solid is directly related to its free energy. If we define the standard state in solution as having an activity of 1, which corresponds to the free energy of zero, then the free energy of transfer from the standard state to Polymorph I is

$$\begin{aligned} G_{\mathrm{I}} - G_{\mathrm{s}} = G_{\mathrm{I}} &= RT\ln a_{\mathrm{I}} - RT\ln 1 = RT\ln a_{\mathrm{I}} \\ &= (H_{\mathrm{I}} - H_{\mathrm{s}}) - T(S_{\mathrm{I}} - S_{\mathrm{s}}) \end{aligned} \tag{2.5}$$

The free energy of transfer from the standard state to Polymorph II is

$$\begin{aligned} G_{\mathrm{II}} - G_{\mathrm{s}} = G_{\mathrm{II}} &= RT\ln a_{\mathrm{II}} - RT\ln 1 = RT\ln a_{\mathrm{II}} \\ &= (H_{\mathrm{II}} - H_{\mathrm{s}}) - T(S_{\mathrm{II}} - S_{\mathrm{s}}) \end{aligned} \tag{2.6}$$

Therefore, the free energy of transfer from Polymorph I to II is

$$\begin{aligned} \Delta G_{\mathrm{I}\rightarrow\mathrm{II}} &= G_{\mathrm{II}} - G_{\mathrm{I}} = (H_{\mathrm{II}} - H_{\mathrm{I}}) - T(S_{\mathrm{II}} - S_{\mathrm{I}}) \\ &= \Delta H_{\mathrm{I}\rightarrow\mathrm{II}} - T\Delta S_{\mathrm{I}\rightarrow\mathrm{II}} \\ &= RT\ln\left(\frac{a_{\mathrm{II}}}{a_{\mathrm{I}}}\right) = RT\ln\left(\frac{f_{\mathrm{II}}}{f_{\mathrm{I}}}\right) \\ &= RT\ln\left(\frac{p_{\mathrm{II}}}{p_{\mathrm{I}}}\right) \approx RT\ln\left(\frac{m_{\mathrm{II}}}{m_{\mathrm{I}}}\right) \end{aligned} \tag{2.7}$$

where:

a is activity,
f is fugacity,
p is vapor pressure,
m is molal concentration.

The approximation is valid if Henry's law is obeyed. Application of the van't Hoff isochore gives

$$\begin{aligned} \frac{d\ln a_{\mathrm{I}}}{d(1/T)} &= \frac{-(H_{\mathrm{I}} - H_{\mathrm{s}})}{R} \approx \frac{d\ln m_{\mathrm{I}}}{d(1/T)} \quad \text{and} \\ \frac{d\ln a_{\mathrm{II}}}{d(1/T)} &= \frac{-(H_{\mathrm{II}} - H_{\mathrm{s}})}{R} \approx \frac{d\ln m_{\mathrm{II}}}{d(1/T)} \end{aligned} \tag{2.8}$$

From which it follows that

$$\begin{aligned} \frac{d\ln\left(\frac{a_{\mathrm{II}}}{a_{\mathrm{I}}}\right)}{d(1/T)} &= \frac{-(H_{\mathrm{II}} - H_{\mathrm{I}})}{R} = \frac{-\Delta H_{\mathrm{I}\rightarrow\mathrm{II}}}{R} \\ &= \frac{d\ln\left(\frac{f_{\mathrm{II}}}{f_{\mathrm{I}}}\right)}{d(1/T)} = \frac{d\ln\left(\frac{p_{\mathrm{II}}}{p_{\mathrm{I}}}\right)}{d(1/T)} \approx \frac{d\ln\left(\frac{m_{\mathrm{II}}}{m_{\mathrm{I}}}\right)}{d(1/T)} \end{aligned} \tag{2.9}$$

When plotting the logarithms of the solubility values against reciprocal absolute temperature, a straight line should result if the solutions behave ideally. The temperature at which the two straight lines for the two polymorphs intersect corresponds to T_t. Alternatively, one could plot the logarithm of the solubility ratios against reciprocal absolute temperature. Again, a straight line should result. Extrapolation of this line to zero indicates the solubility ratio of the two polymorphs is 1, and the temperature corresponds to T_t.

When plotting solubility values against reciprocal temperature, weight-based concentration units (ie, molal

concentrations) are preferred over volume-based (molar) concentration units. This practice will eliminate errors due to thermal expansion of the solutions. However, there is minimum of concern if the solubility ratios are being plotted because the changes in volume cancel each other out.

If dissolution from the solid is solely diffusion controlled, there exists proportionality between the solubility and intrinsic dissolution rate (IDR) by the Noyes–Whitney equation.[31] One could measure the IDR at various temperatures for both polymorphs, analyze the data in the same way as that for solubility, and obtain T_t. This method is particularly helpful when phase transitions occur readily, yet not instantaneously. However, one should exercise caution when comparing dissolution rates of ionic or ionizable compounds in protic solvents because many such dissolution rates have significant contributions from reactive dissolution.

2.3.1.2.1.4 Using solubility/intrinsic dissolution rate and heat of solution data The temperature dependence of solubility is related to the heat of solution, and again for ideal solutions, as shown in Eq. (2.8). If the solubility and the heat of solution are measured at one temperature, the entire solubility-temperature curve can be constructed. Comparing the curves for both polymorphs yields T_t.

Alternatively, one can measure the solubility-IDR ratio at one temperature, which yields the free energy difference at that temperature. The enthalpy difference between the two polymorphs is the difference in the heats of solution (which could also be measured by other means). A simple calculation will yield the T_t.

2.3.1.2.2 Qualitative methods

2.3.1.2.2.1 Using the definition The stability relationship between the two polymorphs at one particular temperature can be determined by one of the following techniques:

1. The higher melting polymorph is more stable at higher temperatures.
2. When spontaneous transition is observed, the resulting polymorph is the more stable phase at the temperature where the transition is observed. The transition could be solid-solid transition isothermally or during heating, or solution-mediated in a nonsolvate-forming solvent.
3. The polymorph with lower solubility in a nonsolvate-forming solvent is more stable.

If the relationships that are determined at two different temperatures are the same, the polymorphic system is monotropic within this temperature range. On the other hand, if the determined relationships are different, then the system is enantiotropic within this temperature range. When an enantiotropic relationship is affirmed, the transition temperature is usually bracketed to a narrower range by determining the stability order at a series of temperatures.

2.3.1.2.2.2 Using the heat of fusion rule To help decide the relationship between two polymorphs, Burger and Ramberger[32,33] developed four thermodynamic rules. One of the two most useful, and reliable, rules is the heat of fusion rule:

> If the higher melting form has the lower heat of fusion, the two forms are usually enantiotropic, otherwise they are monotropic.

Fig. 2.4 illustrates the heat of fusion rule from a thermodynamic point of view. Several precautions should be taken in applying this rule:

1. Be aware of the normal experimental errors in melting points and heats of fusion, especially when they are close for the two polymorphs.
2. Use materials of high crystallinity and high purity. Solids of lower crystallinity usually have lower melting points and heats of fusion than those having high crystallinity. Melting-point depression by impurities is a well-known phenomenon.

2.3.1.2.2.3 Using the heat of transition rule Another useful and reliable rule, developed by Burger and Ramberger,[32,33] is the heat of transition rule for solid-state transitions, which states:

> If an endothermic transition is observed at some temperature, it may be assumed that there is a transition point below it; i.e., the two forms are related enantiotropically. If an exothermic transition is observed at some temperature, it may be assumed that there is no transition point below it; i.e., the two forms are either related monotropically or the transition temperature is higher.

Again, Fig. 2.4 illustrates the thermodynamic point of view. Several precautions should be taken in applying this rule:

1. Be aware that a crystallization exotherm from amorphous phase could be confused with a polymorphic transition. Therefore, materials of high crystallinity should be used in the study. It is a good practice to examine the solid phase after heating beyond the transition to ensure that the exotherm corresponds to a polymorphic transition.
2. One can use the overall heat of transition when the polymorphic transition goes through melting and recrystallization. However, the heat of fusion of the recrystallized polymorph should be comparable to

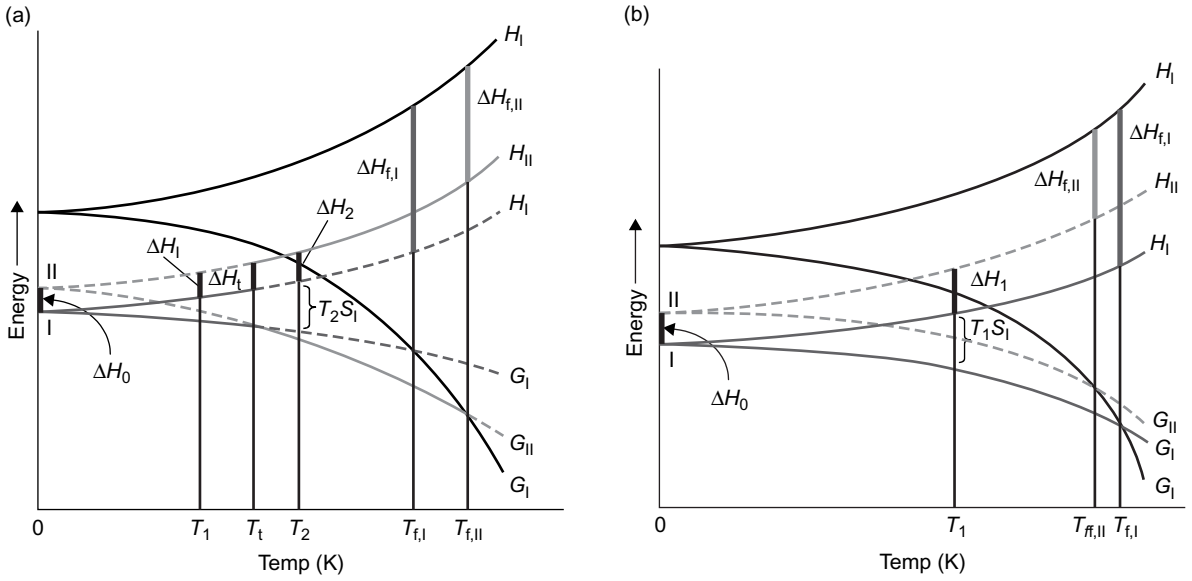


FIGURE 2.4 Schematic illustration of Burger and Ramberger thermodynamic rules. (a) Enantiotropy and (b) Monotropy. *Replotted according to Burger A, Ramberger R. On the polymorphism of pharmaceuticals and other molecular crystals. II. Applicability of thermodynamic rules. Mikrochimi Acta [Wien] 1979;**II**:273–316.*

the reference material to ensure complete crystallization and thus the accuracy of the overall heat of transition.

2.3.2 Solvates/Hydrates

Of the many solvents that are incorporated into the crystal lattice, water is by far the most important for pharmaceutical development. Therefore, all discussions and derivations in this section will use hydrates as examples. However, the principles described are the same and the equations derived can be easily converted and applied to solvates.

2.3.2.1 Anhydrate/Hydrate equilibrium at constant temperature

The equilibrium between the anhydrate and hydrate of a drug (D) at a constant temperature and total pressure can be achieved by equilibrating through the vapor phase. The equilibrium is represented by Eq. (2.10):

$$\mathrm{D}\cdot n\mathrm{H_2O_{(solid)}} \xleftrightarrow{K_d} \mathrm{D_{(solid)}} + n\mathrm{H_2O_{(vapor)}} \quad (2.10)$$

where:

$\mathrm{D_{(solid)}}$ is the anhydrate solid,
$\mathrm{D}\cdot n\mathrm{H_2O_{(solid)}}$ is the hydrate solid,
n is the stoichiometry of the hydrate,
K_d is the hydration/dehydration equilibrium constant between the two solids.

The equilibrium constant can be described by Eq. (2.11) and then simplified based on the assumption that water vapor acts as an ideal gas, and the activities of the solids are taken to be unity:

$$K_d^{\text{vapor}} = \frac{\left[a_{\mathrm{D_{(solid)}}}\right]\left[a^c_{\mathrm{H_2O_{(vapor)}}}\right]^n}{\left[a_{\mathrm{D}\cdot n\mathrm{H_2O_{(solid)}}}\right]} = \left[a^c_{\mathrm{H_2O_{(vapor)}}}\right]^n = [RH^c]^n = \left[\frac{p^c}{p^s}\right]^n \quad (2.11)$$

In this equation, a is the activity, superscript c represents critical values, RH represents relative humidity, p^c represents the critical water vapor pressure, and p^s represents the saturated water vapor pressure at this temperature.

Three situations exist:

1. $p > p^c$; $RH > RH^c$; $a_{\mathrm{H_2O_{(vapor)}}} > a^c_{\mathrm{H_2O_{(vapor)}}}$: The hydrate is more stable than the anhydrate; therefore, the anhydrate will spontaneously convert to the hydrate thermodynamically.
2. $p < p^c$; $RH < RH^c$; $a_{\mathrm{H_2O_{(vapor)}}} < a^c_{\mathrm{H_2O_{(vapor)}}}$: The anhydrate is more stable than the hydrate; therefore, the hydrate will spontaneously convert to the anhydrate thermodynamically.
3. $p = p^c$; $RH = RH^c$; $a_{\mathrm{H_2O_{(vapor)}}} = a^c_{\mathrm{H_2O_{(vapor)}}}$: The hydrate and the anhydrate have equal stability, and they will coexist thermodynamically. There will be no transformation in either direction.

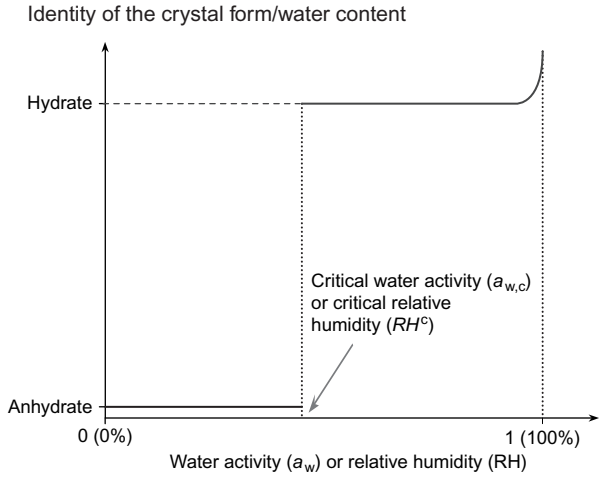


FIGURE 2.5 Thermodynamic phase diagram of an anhydrate/hydrate pair.

A typical phase diagram for anhydrate/hydrate is shown in Fig. 2.5. The curved portion, at high RH, represents deliquescence and further dilution of the solution. Anhydrate/hydrate equilibration through the vapor phase is a slow process, and it can take weeks to months to achieve equilibrium. Therefore, studies are usually conducted to take advantage of the higher mobility in the solution state to attain equilibrium within a short period of time.[34,35] In these studies, water-miscible organic solvents were employed to systematically modify the water activities in these solutions. Solids of anhydrate or hydrate, or mixtures of the two, were suspended in a series of these solutions and equilibrated at a defined temperature with agitation. The solids were recovered after certain periods of time, which were much shorter than the time required for equilibration through the vapor phase, and examined by appropriate analytical techniques. The stability order at each individual water activity value was determined by the nature of the solid recovered. The critical water activity was then bracketed to a narrow range. Once developed, this method has been applied successfully to many other anhydrate/hydrate systems.[36–41]

When an organic/water solution is in equilibrium with its vapor phase, the water activities in the solution and the vapor phases are the same. Therefore, it is common to further equate the water activity in solution to the RH in the vapor phase. For example, 0.35 water activity is equivalent to 35% RH. The critical water activity values obtained from solution equilibration for anhydrate/hydrate systems are also equated to the critical RH values obtained from the vapor phase equilibration (RH^c). The reason why this can be done is demonstrated next.

In the case of solution phase equilibration, water molecules released via dehydration are not present as water vapor; instead, they simply become part of the solution. Therefore, the equilibrium of interest and the corresponding equilibrium constant are shown in Eqs. (2.12) and (2.13), respectively:

$$\mathrm{D}\cdot n\mathrm{H_2O_{(solid)}} \xleftrightarrow{K_d} \mathrm{D_{(solid)}} + n\mathrm{H_2O_{(solution)}} \tag{2.12}$$

$$K_d^{\text{solution}} = \frac{\left[a_{\mathrm{D_{(solid)}}}\right]\left[a^c_{\mathrm{H_2O_{(solution)}}}\right]^n}{\left[a_{\mathrm{D}\cdot n\mathrm{H_2O_{(solid)}}}\right]} = \left[a^c_{\mathrm{H_2O_{(solution)}}}\right]^n \tag{2.13}$$

This equilibrium can be converted to the Eq. (2.10) type of equilibrium (ie, water as vapor) through a further equilibrium of solution-vapor under constant water activity (Eq. 2.14). However, the process of vaporization under constant water activity is at equilibrium, having ΔG of zero and K_{vap} of unity:

$$n\mathrm{H_2O_{(solution)}} \xleftrightarrow{K_{vap}} n\mathrm{H_2O_{(vapor)}} \tag{2.14}$$

The addition of Eqs. (2.12) and (2.14) affords Eq. (2.10). Therefore, the equilibrium constant in Eq. (2.10) is the product of those in Eqs. (2.12) and (2.14) (Eq. 2.15):

$$K_d^{\text{vapor}} = K_{vap}\cdot K_d^{\text{solution}} = K_d^{\text{solution}} \tag{2.15}$$

Thus, the equilibrium constant obtained by equilibrating through the vapor phase is equal to that obtained through equilibrating through the solution phase. In the following section, superscripts and subscripts will not be used to differentiate the different routes of determination.

2.3.2.2 *Temperature dependence of anhydrate/hydrate equilibrium*[42]

When temperature varies, the van't Hoff equation (also known as the *van't Hoff isochore*), which describes the temperature dependence of any equilibrium constant, can be applied. This affords the determination of the temperature dependence of the critical water activity or RH ($a^c_{H_2O}$ or RH^c), and the critical water vapor pressure (p^c), thus establishing the complete phase diagram with regard to both RH (or water vapor pressure) and temperature. Applying the van't Hoff equation to the anhydrate/hydrate equilibrium, the following equation is obtained:

$$\frac{d\ln K_d}{dT} = \frac{nd\ln(a^c_{H_2O})}{dT} = \frac{nd\ln(RH^c)}{dT} = \frac{nd\ln\left(\frac{p^c}{p^s}\right)}{dT} = \frac{\Delta H_d}{RT^2} \tag{2.16}$$

where ΔH_d is the dehydration enthalpy and R is the gas constant. Rearranging Eq. (2.16) affords the following equation:

$$d\ln K_d = d\ln(a^c_{H_2O}) = d\ln(RH^c) = d\ln\left(\frac{p^c}{p^s}\right) = -\frac{\Delta H_d}{nR} d\ln\left(\frac{1}{T}\right) \quad (2.17)$$

Assuming a constant dehydration enthalpy within the temperature range studied, the definite integral of this differential equation (Eq. 2.17) between temperatures T_1 and T_2 is given by the following equation:

$$\frac{\Delta H_d}{nR}\left(\frac{1}{T_1}-\frac{1}{T_2}\right) = \ln K_d^2 - \ln K_d^1 = \ln(a^c_{H_2O})_2 - \ln(a^c_{H_2O})_1 = \ln(RH^c)_2 - \ln(RH^c)_1 = \ln\left(\frac{p^c}{p^s}\right)_2 - \ln\left(\frac{p^c}{p^s}\right)_1 \quad (2.18)$$

Since $\Delta G_d = \Delta H_d - T\Delta S_d$ and $\Delta G_d = -RT\ln K_d$, it follows that

$$\ln K_d = \ln(a^c_{H_2O}) = \ln(RH^c) = \ln\left(\frac{p^c}{p^s}\right) = -\frac{\Delta H_d}{nRT} + \frac{\Delta S_d}{nR} \quad (2.19)$$

Here ΔS_d is the dehydration entropy. Therefore, a plot of the natural logarithm of K_d or $a^c_{H_2O}$, or RH^c, or p^c/p^s, versus the reciprocal temperature gives a straight line. The slope of the line is equal to $-\Delta H_d/nR$ and the intercept is equal to $\Delta S_d/nR$. Based on this set of equations, and on the fact that dehydration is endothermic, it is apparent that the critical water activity or critical RH increases upon raising the temperature. Therefore, it is desirable to keep hydrates in a cool, humid environment to avoid dehydration.

Another important parameter of the anhydrate/hydrate system is the critical temperature, T_c, which is defined as the temperature at which the critical water activity is unity or critical RH is 100%. In other words, it is the higher limit of the temperature range where a hydrate is thermodynamically stable at some water activity or RH. At temperatures higher than T_c, hydrate cannot be the stable phase. Applying this to Eq. (2.18) affords this equation:

$$\ln K_d = \ln(a^c_{H_2O}) = \ln(RH^c) = \ln\left(\frac{p^c}{p^s}\right) = \frac{\Delta H_d}{nRT_c} - \frac{\Delta H_d}{nRT} \quad (2.20)$$

Comparing Eqs. (2.19) and (2.20), it is apparent that $T_c = \Delta H_d/\Delta S_d$. Thus, T_c is characteristic of the anhydrate/hydrate system of interest.

In studying the physical stability of hydrates, p^c is frequently used. Therefore, it is useful to extend this analysis beyond $K_d, a^c_{H_2O}, RH^c$, and p^c/p^s. Rearranging Eq. (2.18) affords the following:

$$\ln\frac{(p^c)_2}{(p^c)_1} - \ln\frac{(p^s)_2}{(p^s)_1} = \frac{\Delta H_d}{nR}\left(\frac{1}{T_1}-\frac{1}{T_2}\right) \quad (2.21)$$

The temperature dependence of saturated water vapor pressure can be described by the well-known Clausius–Clapeyron equation:

$$\ln\frac{(p^s)_2}{(p^s)_1} = \frac{\Delta H_{vap}}{R}\left(\frac{1}{T_1}-\frac{1}{T_2}\right) \quad (2.22)$$

Here, ΔH_{vap} is the enthalpy of vaporization of water at the temperature of interest. Coupling Eqs. (2.21) and (2.22) yields

$$\ln\frac{(p^c)_2}{(p^c)_1} = \left(\frac{\Delta H_d}{nR} + \frac{\Delta H_{vap}}{R}\right)\left(\frac{1}{T_1}-\frac{1}{T_2}\right) \quad (2.23)$$

Eq. (2.23) dictates that the plot of $\ln p^c$ against the reciprocal temperature will exhibit a greater slope than that of $\ln K_d (\Delta H_d/nR + \Delta H_{vap}/R$ versus $\Delta H_d/nR)$. This phenomenon was previously observed, but it has not been explained.[43]

2.3.3 Cocrystals

Since cocrystals differ from solvates only by the physical state of the pure components, the stability treatment is expected to be analogous to that used for solvates. Indeed, an extension of the stability treatment was recently applied to cocrystals[44] and is discussed in the following sections. The equilibrium among the drug (D), the cocrystal former (CCF), and the corresponding cocrystal (CC; ie, D·nCCF) are shown in Eq. (2.24). In this case, the CCF is assumed to come into intimate contact with the drug through a solution; however, the treatment is valid if the contact is achieved by other routes:

$$D_{solid} + nCCF_{solution} \xleftrightarrow{K_c} CC_{solid} \quad (2.24)$$

Here, n is the stoichiometry of the cocrystal and K_c is the cocrystal formation equilibrium constant. Analogous to the anhydrate/hydrate or nonsolvate/solvate system, K_c can be expressed by the activities of the species involved in the equilibrium (Eq. 2.25) and simplified based on the assumption that CCF

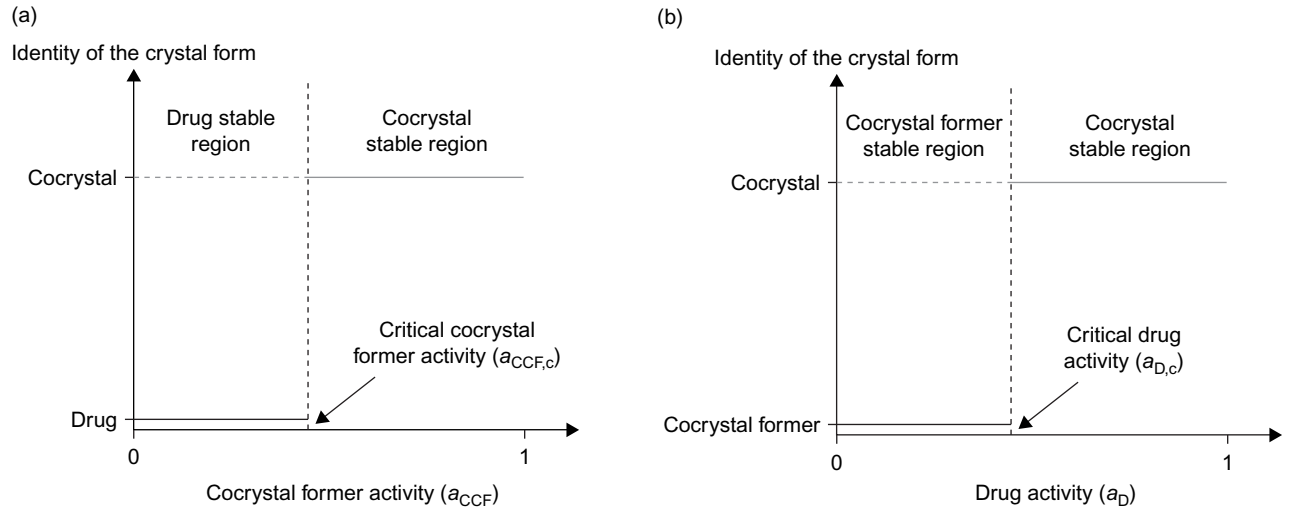


FIGURE 2.6 Thermodynamic phase diagrams of cocrystals. (a) With respect to the activity of the cocrystal former and (b) With respect to the activity of the drug.

molecules behave ideally in the solution, and the activities of the solids are taken to be unity:

$$K_c = \frac{\left[a_{CC_{(solid)}}\right]}{\left[a_{D_{(solid)}}\right]\left[a^c_{CCF_{(solution)}}\right]^n} = \left[a^c_{CCF_{(solution)}}\right]^{-n} \qquad (2.25)$$

Here, $a^c_{CCF_{(solution)}}$ is the critical CCF activity in solution (and can be expressed as $a_{CCF,c}$). As a first approximation, the activity of CCF in the solution equals the ratio of the concentration of the free CCF (uncomplexed with D) in solution, and the solubility of CCF in the absence of D in the same solvent. Again, three situations exist:

1. $\left[a_{CCF_{(solution)}}\right] < \left[a^c_{CCF_{(solution)}}\right]$: D is favored over CC; therefore, CC will spontaneously decompose to D thermodynamically.
2. $\left[a_{CCF_{(solution)}}\right] > \left[a^c_{CCF_{(solution)}}\right]$: CC is favored over D; therefore, CC will spontaneously form from D and CCF thermodynamically.
3. $\left[a_{CCF_{(solution)}}\right] = \left[a^c_{CCF_{(solution)}}\right]$: D and CC have equal stability, and they will coexist in a suspension of the CCF solution.

The same equations can be written if the drug is in solution and CCF and CC are in the solid state. This thermodynamic treatment can be easily extended to cocrystal systems with multiple stoichiometries. The phase diagrams are shown in Fig. 2.6.

2.3.4 Amorphous solids

If a crystalline solid is heated beyond its melting temperature and cooled back to its original temperature, and if sufficient time is not allowed for nucleation to occur, the liquid will become amorphous at temperatures below its melting point. That is, it will lose all long-range order. Initially, the supercooled liquid was able to adjust itself to the changing environment (temperature), giving off excess enthalpy, entropy, and volume, following the equilibrium line extended from the liquid state (Fig. 2.7). The amorphous material in this temperature range is called a *rubber* or a *supercooled liquid*. However, at a certain lower temperature, due to the increased viscosity and decreased molecular mobility, the amorphous material will not be able to continue to follow the equilibrium line extended from the liquid state. Instead, it will deviate from the line, quenching the excess thermal motions that accommodate excess quantities such as free energy, enthalpy, entropy, and volume. The amorphous phase in this temperature range is called a *glass*. This transition of states is called a *glass transition*, and the temperature at which this transition occurs is called the *glass transition temperature* (T_g). The glassy state is a kinetically frozen state and is not thermodynamically stable. Therefore, a glass will relax gradually over time, releasing excess enthalpy, entropy, and volume to the environment to achieve a more stable state. This process is termed *physical aging* or *annealing* (discussed further in Section 2.7.1).

Theoretically, one could continue to extrapolate the liquid line beyond T_g. In doing so, it will eventually intersect the crystalline line and pass beyond it. This presents a conceptual contradiction, called the *Kauzmann paradox*,[45] which dictates that the disordered supercooled liquid has less entropy than the ordered crystal. The temperature at which this occurs is called the *Kauzmann temperature* (T_K). In reality, this entropy

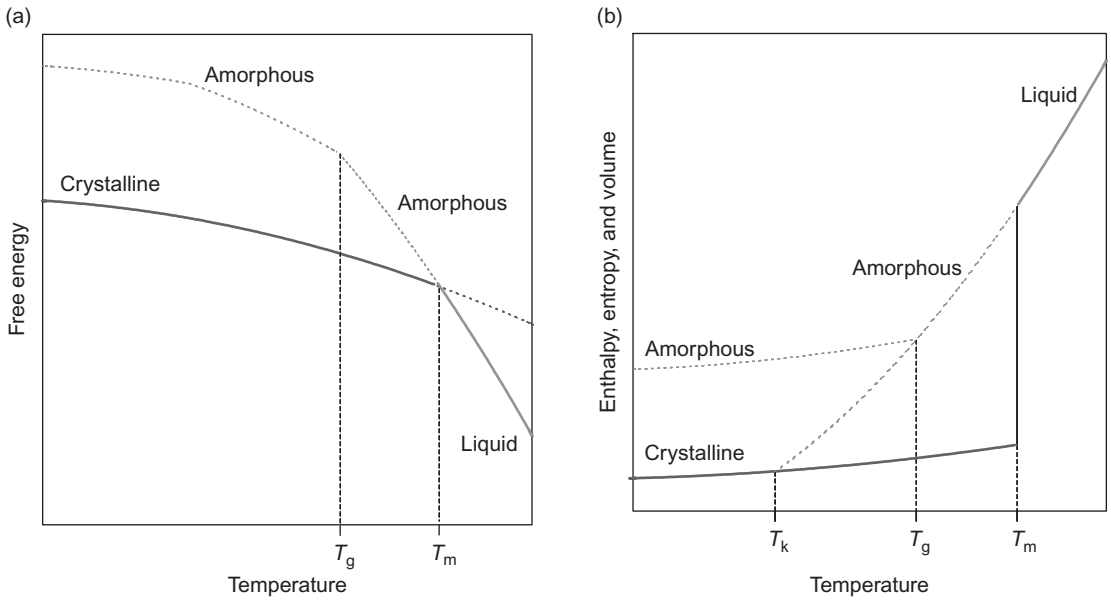


FIGURE 2.7 Phase diagrams of amorphous solids. (a) Free energy and (b) Enthalpy, entropy, and volume.

crisis is avoided through the intervening glass transition. Compared to its crystalline counterpart, the amorphous solid possesses excess thermodynamic quantities, including free energy, enthalpy, entropy, and volume (Fig. 2.7). These excess quantities are termed *configurational* quantities. From the phase diagram (Fig. 2.7), it is clear that, unlike polymorphs, solvates, and cocrystals, amorphous phases are always metastable with respect to the crystalline phases, regardless of the environment (temperature, pressure, etc.). Unlike the first-order phase transitions, such as melting or vaporization, no discontinuity exists at the glass transition for thermodynamic properties, such as volume, enthalpy, and entropy. However, the isobaric expansivity, α, isothermal compressibility, κ, and isobaric heat capacity, C_p, all show a step change around T_g, indicating a second-order phase transition according to the definition by Ehrenfest.[46]

The amorphous phase presents unique characteristics and challenges to pharmaceutical development, and this topic is discussed in detail in Section 2.7.

2.4 PHARMACEUTICAL RELEVANCE AND IMPLICATIONS

As mentioned at the beginning of this chapter, crystalline solids are especially relevant to pharmaceutical development. Regardless of the types of formulations, APIs are mostly crystalline. The majority of APIs are polymorphs of the parents and salts.[47] A significant proportion of APIs are hydrates.[47] Solvates with various organic solvents are not usually used as APIs. However, they are frequently encountered in crystallization and characterization of pharmaceutical solids. Cocrystals, a reemerging class of solids, have yet to be used intentionally in pharmaceutical products. However, they offer the potential to engineer properties of solids that are critical to pharmaceutical development. Although crystalline APIs are usually preferred during pharmaceutical development, varying degrees of amorphous content can be generated inadvertently during pharmaceutical processing, such as milling, wet granulation, drying, and compaction (see chapter: Drug Stability and Degradation Studies). Many pharmaceutical excipients, such as naturally existing polymers (eg, cellulose and starch), synthetic polymers (eg, poly(methyl methacrylate), poly(acrylic acid), and poly(glycolic acid)), and other chemically modified natural or synthetic polymers, are either amorphous or partially amorphous. Small amounts of amorphous content exist in almost every crystalline API. The presence of amorphous content in the API itself or other components of a drug product may have confounding impacts on the physicochemical properties and ultimate performance of the drug product.

Owing to the differences in the internal structure (composition, dimensions, shape, symmetry, capacity (number of molecules), and void volumes of their unit cells for crystalline phases, and the lack of periodicity for the amorphous phases), various solids exhibit a variety of properties. Over the years, these properties have been summarized for polymorphs,[48–51] but they are properly applicable to all solids. These properties are listed in Table 2.2.

TABLE 2.2 Physical Properties That Differ Among Various Solids[51]

1. Packing properties
   - a. Molar volume and density
   - b. Refractive index
   - c. Conductivity, electrical, and thermal
   - d. Hygroscopicity
2. Thermodynamic properties
   - a. Melting and sublimation temperatures
   - b. Internal energy (ie, structural energy)
   - c. Enthalpy (ie, heat content)
   - d. Heat capacity
   - e. Entropy
   - f. Free energy and chemical potential
   - g. Thermodynamic activity
   - h. Vapor pressure
   - i. Solubility
3. Spectroscopic properties
   - a. Electronic transitions (ie, ultraviolet—visible absorption spectra)
   - b. Vibrational transitions (ie, infrared absorption spectra and Raman spectra)
   - c. Rotational transitions (ie, far infrared or microwave absorption spectra)
   - d. Nuclear spin transitions (ie, nuclear resonance spectra)
4. Kinetic properties
   - a. Dissolution rate
   - b. Rates of solid-state reactions
   - c. Stability
5. Surface properties
   - a. Surface free energy
   - b. Interfacial tensions
   - c. Habit (ie, shape)
6. Mechanical properties
   - a. Hardness
   - b. Tensile strength
   - c. Compactibility, tableting
   - d. Handling, flow, and blending

Many properties that differ among crystal forms provide the means to detect them, such as melting points and spectroscopic properties. Some properties are particularly important for pharmaceutical development, having a significant impact on the performance and quality of drug products. They will be discussed in more detail in the following sections.

2.4.1 Solubility

As shown in Section 2.3.1, the solubility values of different polymorphs are directly related to their free energy. Therefore, different polymorphs will have different solubilities in the same solvent, and the difference in solubility will be a function of temperature (Fig. 2.8). At one temperature (except for T_t), the metastable polymorph will be more soluble than the stable polymorph. It is this difference in solubility that is causing the difference in dissolution rates and driving the solution-mediated phase transformation (SMPT) (Section 2.5.3) for polymorphs. A survey of the solubility differences between polymorphic pairs has been conducted ($n = 81$).[52] The general trend reveals that the ratio of polymorph solubility is typically less than 2, although occasionally higher ratios can be observed.[52] Based on thermodynamics, the solubility ratio of the polymorphs is independent of the nature of the solvents, provided that Henry's law is obeyed (Section 2.3.1). In most dilute solutions, this is a valid statement.

Adducts (solvates, hydrates, and cocrystals), on the other hand, exhibit different solubility behaviors. The solubility ratio, even the rank order, changes in solvents with different activity values. As in the case of polymorphs, this is also driven by thermodynamics. In pure organic solvents, where the activity of the solvent is nearly 1, the solvate with the specific solvent will be the most stable phase and have the lowest solubility. Other solid phases, such as polymorphs, hydrates, amorphous phases, and solvates with other organic solvents, will be metastable and have higher solubility values. Hydrates, as a special case of solvates, will have the same characteristics. In the same survey ($n = 24$),[52] a similar trend is observed for anhydrate/hydrate solubility ratios, although anhydrate/hydrate solubility ratios appear to be more spread out and higher than the typical ratio for polymorphs. The solubility of cocrystals is a function of the concentration of the cocrystal former.[53,54] The dependence is very much analogous to the solubility product (K_{sp}) for insoluble salts.

Thermodynamic analysis of amorphous phases indicates that it possesses higher free energy than its crystalline counterpart at all temperatures below the melting point. Therefore, amorphous materials will have a higher solubility than crystalline solids at all temperatures (below the melting point) in any solvent.[55–59] This is of special interest to the delivery of poorly soluble drug molecules and oral bioavailability enhancement.

2.4.2 Dissolution rate and bioavailability

As mentioned earlier, if dissolution from the solid is solely diffusion-controlled, proportionality exists between the solubility and dissolution rate as determined by the Noyes–Whitney equation.[31] Thus, the differences in solubility translate directly to the differences in IDR among various solids. When particle size and distribution are similar, they also translate to the powder dissolution rates. Some of the classic examples include dissolution of the solid

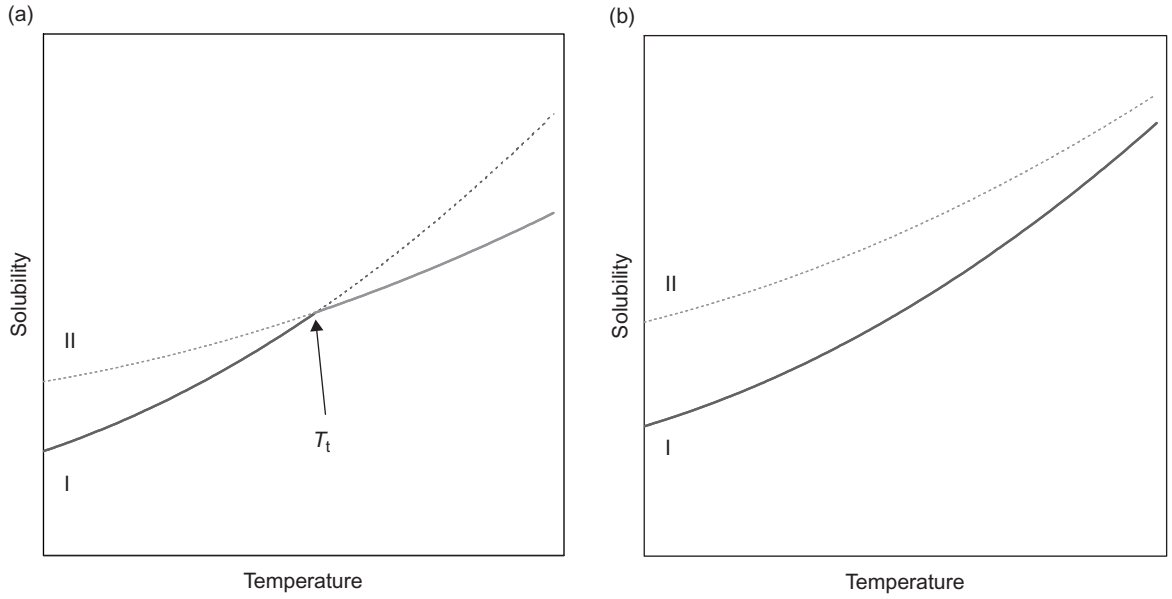


FIGURE 2.8 Solubility values of polymorphs as a function of temperature. (a) Enantiotropy, (b) Monotropy.

forms of succinylsulfathiazole[60] and iopanoic acid.[61] As generally understood, oral absorption is determined by the rate of dissolution of a dosage form in the gastrointestinal (GI) tract and the rate of drug molecule permeation through the GI membrane. Drug absorption can be improved by rendering the drug amorphous for biopharmaceutical classification system (BCS) class II and IV compounds, where low solubility presents a significant barrier for oral absorption. Therefore, utilizing metastable solid forms, such as amorphous phases, and amorphous solid dispersions[62–69] is a powerful tool in combating the poor absorption profiles of many modern drug candidates. Selecting the appropriate solid forms for development is thus critical to the facile development of high quality products.

One important consideration in this approach is that this solubility and dissolution enhancement is a transient effect. The higher concentrations achieved in the in vitro or in vivo dissolution media, beyond the equilibrium solubility, create driving forces for the nucleation and crystallization of the stable phases. Eventually, the stable phases will nucleate, and crystallize. However, a certain amount of time is required before the depletion of concentrations in solution. During this lag time, significant absorption could be accomplished. For some compounds, extremely high permeability leads to the effective transport of the molecules through the GI membrane. In these cases, the concentration buildup in the GI fluid will not occur.

2.4.3 Hygroscopicity

Moisture interacts with solids via four primary modes: adsorption, absorption, deliquescence, and lattice incorporation.[70,71] Water molecules can be adsorbed onto a solid surface by interacting with the molecules on the surface. Since water is a polar molecule capable of forming hydrogen bonding, polar moieties on the solid surface are important factors governing the affinity toward water adsorption. Particle size plays a role in moisture uptake by affecting the available surface area. Under normal conditions, water adsorption leads to monolayer formation, as well as a few additional molecular layers at higher RH values. The solid-water interactions are the strongest between the surface and water molecules in the monolayer, and they become weaker as the distance from the surface increases. Therefore, water molecules in the adsorbed layers beyond the monolayer better resemble bulk water, which may bear some importance to chemical instability.

A more significant mode of water–solid interactions is perhaps the absorption or sorption, which refers to the water molecules' penetration into the bulk solid, incorporation into the defects or amorphous regions, and formation of a solution. Due to the high free energy nature of the amorphous form, its affinity toward water molecules is much greater, resulting in much elevated moisture sorption. It has been shown that moisture adsorption can account for only a very insignificant portion of moisture content typically encountered in

pharmaceutical solids (eg, a few percent), for which the sorption mechanism is primarily responsible.[72] *Deliquescence* refers to the formation of a saturated solution around the solids when exceeding its critical RH. This mode of water–solid interaction is only important for relatively water-soluble compounds at high RH. Hydrates form when water molecules are incorporated into the crystal lattice. In this case, a critical water activity or RH exists that governs the physical stability of the respective hydrate form (Section 2.3.2).

Moisture uptake can be greatly affected by solid forms. Adsorption is a surface property; therefore, the surface of a solid form is a determining factor. In that sense, a more stable polymorph may not necessarily adsorb the least amount of moisture; rather, what matters are the moieties that are exposed on the surface. However, the amorphous content or defects are the primary cause of the moisture uptake normally seen with pharmaceuticals, as discussed previously, due to its high energetic nature and a different mode of interaction. A small amount of amorphous content virtually exists for every crystalline API and is likely exacerbated by API treatments such as milling and by formulation processing such as wet granulation and drying. Therefore, its influence on moisture uptake and distribution, as well as subsequent product attributes, could be enormous.

2.4.4 Reactivity and chemical stability

Solid forms can affect the reactivity and chemical stability of both API and drug in the finished product. The way in which solid forms influence their chemical stability is more apparent by examining the mechanisms of "solid-state" reactions.

2.4.4.1 Topochemical reactions

The reactivity of a solid in a topochemical (or true solid-state) reaction is completely determined by its crystal lattice. This type of reaction is facilitated by the proper positioning and orientation of the reacting moieties in the crystal lattices, so it requires minimum movement of the reacting molecules. Molecular packing in the crystal lattice is the primary factor dictating the rate of a topochemical reaction.

Solid forms are the best manifestation of topochemical reactions based on their very nature. Solid forms can determine the end-reaction product. For example, 2-ethoxycinnamic acid showed polymorph-dependent photodimerization products.[73] The α polymorph formed a cis-dimer, the β polymorph formed a trans-dimer, while dimerization was not observed in solution. Solid forms can also determine whether a reaction can occur. For example, *p*-methylcinnamic acid undergoes photodimerization only with the β polymorph.[74] Its α polymorph is not reactive because the distance separating the neighboring cinnamic acid double bonds is too long to allow a (2 + 2) cyclization. For prednisolone and hydrocortisones, the reactivity toward O_2 was observed only in the hexagonal solvates, where solvent channels were speculated to allow oxygen to access host molecules.[75,76] Increasing molecular mobility by destroying a crystalline order can actually slow down a topochemical reaction. For example, the rearrangement of methyl *p*-dimethylaminobezenesulfonate proceeded 25 times faster in solid state than in the melt.[77,78]

2.4.4.2 Nontopochemical reactions

The degradation of most pharmaceuticals is not topochemical in nature. These reactions require significant movement of the reacting molecules and are not true solid-state reactions. Nevertheless, these reactions are modulated by the physicochemical and particle characteristics of the solids, such as crystallinity, moisture content, solubility, pH of the aqueous surface layer, particle size, and melting temperature, although it may be difficult to deconvolute the respective contributions.

Chemical stability of most pharmaceuticals is modulated by moisture. The earlier model proposed by Leeson and Mattocks[79] sheds some light on drug degradation, where a film of saturated solution was assumed to form around solid particles, although deliquescence is not generally expected below critical RH. The concepts of tightly bound (eg, monolayer) and loosely bound (eg, those beyond monolayer) moisture may also shed some light, despite their oversimplified nature. Probably the most significant is the moisture sorption into amorphous regions or defects, leading to significant increase in moisture uptake, as discussed previously.

Besides the topochemical reactions where crystal lattice dictates reactivity, solid forms can affect chemical stability by modifying the particular characteristics. First, the moieties exposed on crystal surfaces could vary significantly among different solid forms (or morphologies), which could either accelerate or inhibit a particular reaction. For example, the reactivity of flufenamic acid polymorphs with ammonia was related to accessibility of ammonium to host molecules on major faces, but not to the relative energetics of the polymorphs.[80] Second, the solid forms can differ in their aqueous solubility, as mentioned previously, which is certainly important for degradation through the Leeson–Mattocks model. Third, moisture sorption could be drastically more significant for certain solid forms. Amorphous content has a higher affinity to water molecules and forms solutions with the latter.

Since water is a great plasticizer, the water molecules absorbed into the amorphous regions can cause a significant increase in molecular mobility and are potentially harmful to chemical stability.[72,81] Certain higher energetic solid forms that may be used in pharmaceutical development, such as "desolvated solvate," tend to absorb more moisture and cause greater chemical degradation. Complete or partial phase transformations can occur inadvertently during manufacturing, which can profoundly change chemical degradation in drug products.[82]

2.4.5 Mechanical properties

In different solids, the molecules are arranged differently. The nature and extent of interactions among the molecules also varies among different solids. Intuitively, one would anticipate that these solids would respond differently under mechanical stresses. Under compression, plastic flow along slip planes has been well documented as a mechanism for consolidation.[83] In the case of acetaminophen polymorphs, the stable polymorph (Form I, monoclinic) was found to be poorly compressible, requiring a significant amount of excipients to form tablets.[84–86] On the other hand, metastable polymorphs (Form II, orthorhombic) have been reported to have a higher cohesion index[87] and thus improved tableting properties. Careful analysis of crystal structure and face indexing[88] revealed that these crystals cleaved parallel to the {001} plane. This cleavage occurred due to a well-developed slip system, which is attributed to the presence of two-dimensional molecular sheets in Form II solids. These improved tableting properties of Form II were also confirmed by Heckel analysis, compressibility, and compactibility assessments.[89] Similarly, sulfamerazine metastable polymorph I has well-defined slip planes in its crystal lattice, thus exhibiting superior plasticity and therefore greater compressibility and tabletability than the stable Polymorph II.[90] Water produced by the crystallization in *p*-hydroxybenzoic acid monohydrate facilitates plastic deformation of the crystals, thereby enhancing their bonding strength and forming stronger tablets than the anhydrate.[91]

2.5 TRANSFORMATIONS AMONG SOLIDS

As discussed in Section 2.3.1, polymorphic transformation is widely used in understanding polymorphic relationships. This section focuses on the various pathways for transformation between crystalline solid forms. Crystallization from amorphous materials, owing to its importance to pharmaceutical development, will be discussed separately in Section 2.7. In general, thermodynamics dictates the possibility of transformation between the various phases. Therefore, the discussion in Section 2.3 and an understanding of the particular system should be kept in mind.

Most of the transitions described here are second-order transitions; that is, they are kinetic in nature (except for melting). Therefore, the timescale of experiments can drastically change the pathways that a solid will take in transforming into the other forms. Sometimes the resulting solid phase can also be different.

2.5.1 Induced by heat

Molecular motions and mobility in solids generally increase when the temperature of the solid is raised. In addition, the free energy of the various phases changes with respect to temperature, as determined by thermodynamics. Therefore, heating is particularly effective in inducing phase transformations.

2.5.1.1 Polymorphic transitions

As shown in Section 2.3.1 earlier in this chapter, a pair of polymorphs can be related to each other, either monotropically or enantiotropically. When heated, they exhibit different behaviors upon undergoing transformations. Fig. 2.9 illustrates all potential polymorphic transitions that can be observed when a polymorphic solid is heated in differential scanning calorimetry (DSC). The descriptions of the phase transitions are placed next to the thermal events in DSC traces and are explained in detail in Table 2.3. This table is arranged according to the polymorphic system, then the starting polymorph, and then the different possibilities of the transitions. Not all situations constitute polymorphic transition. When examining the thermal traces for polymorphic transitions, it is helpful to have the phase diagrams on hand (Fig. 2.2).

Most of the transitions listed in Table 2.3 have been observed in real life, except for E II-3. In this case, Polymorph II converts via a solid-solid route to Polymorph I at a temperature below T_t. Unlike in E II-2, where the resulting Polymorph I converts back to Polymorph II when heated past T_t, Polymorph I melts without converting to Polymorph II. The reason that E II-3 is not observed is that the mobility of the solid is higher at higher temperatures; therefore, the kinetic resistance to polymorphic transitions is also likely to be less at higher temperatures.[50] For one pair of enantiotropic polymorphs, if the conversion in one direction is observed at lower temperatures, the likelihood of not

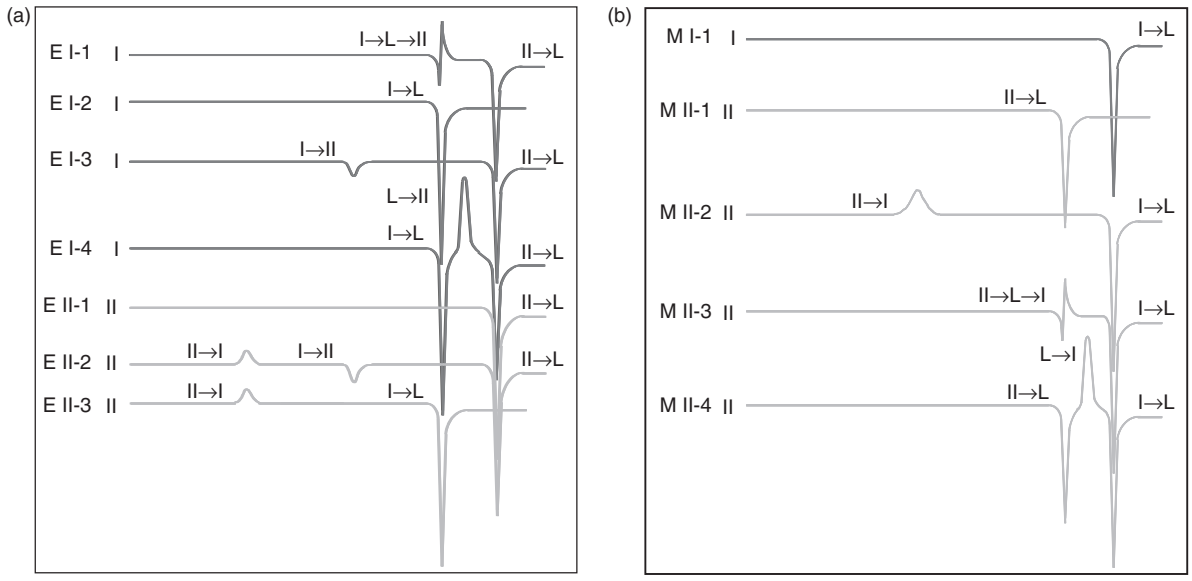


FIGURE 2.9 Potential polymorphic transitions induced by heat. (a) Enantiotropy and (b) Monotropy.

observing the conversion in the other direction at higher temperatures is very low.

Experimentally, one could heat the solids to a temperature slightly above the transition, cool the solid back to ambient, reheat, and observe the reversibility of the transitions. Techniques like this have been used extensively to demonstrate the enantiotropic relationship between two polymorphs.[92–94] Due to differences in the mobility of solids at different temperatures, one can only draw positive conclusions, not negative ones. In other words, if reversibility of the transition is observed, it is conclusive for an enantiotropic relationship. However, when reversibility is experimentally not observed, one cannot conclude that the polymorphs are monotropically related.

For purposes of illustration, only dimorphic systems are described here. In reality, many polymorphic systems have three or more polymorphs. Therefore, the transformations upon heating could be much more complex than those demonstrated here. In complex cases, several complementary analytical methods should be used together in order to understand the system.

2.5.1.2 *Dehydration/Desolvation*

The thermodynamic of solvates and hydrates (Section 2.3.2) dictates that they will be less stable at higher temperatures. Indeed, heating has been routinely applied to remove crystallization solvents that are incorporated into the lattice (ie, desolvation). Thermogravimetric analysis (TGA) is one of the main characterization methods for solvates/hydrates. In principle, TGA monitors the weight change as a function of temperature. The total weight loss corresponds to the stoichiometry of the solvate/hydrate, while the temperature at which the weight loss begins is indicative of the physical stability of the solvate/hydrate.

The resulting solvent-free solids could be different when desolvating different solvates, desolvating the same solvate under different conditions, or desolvating the same solvate under the same conditions but different batches, and with different amounts of residue solvent.[95] Therefore, desolvation is one of the major approaches in polymorph screening to search for polymorphs.

2.5.1.3 *Cocrystal formation*

Before the term *cocrystal* was invented, this type of adducts was called *solid-state molecular complexes*. Many of them have been reported in the pharmaceutical literature. Very often, the discoveries resulted from experiments to construct binary melting-point phase diagrams. In the early days, thermomicroscopy was used. One would watch the solids as they were heated. Very often, after eutectic melting, the liquid would crystallize into a new solid, which might correspond to a solid-state complex. Later, the construction of phase diagrams shifted to DTA/DSC when they became widely available. The common practice was to melt the mixtures of solids first, cool them, and let them crystallize to ensure perfect thermal contact between the sample and the pans and intimate mixing of the two components, and then collect melting points during the second heating. This practice led to an even higher probability of the crystallization of complexes because the liquid spent more time at higher temperatures and had more opportunity for crystallization.

TABLE 2.3 Potential Thermal Events Observed When Heating a Polymorphic Solid

Thermal trace ID	Polymorphic system	Starting polymorph	Transitions (In order of increasing temperature)
E I-1	Enantiotropic	I	Melting of I with concurrent recrystallization of II
			Melting of II
E I-2	Enantiotropic	I	Melting of I
E I-3	Enantiotropic	I	Solid-state conversion from I to II
			Melting of II
E I-4	Enantiotropic	I	Melting of I
			Recrystallization of II from liquid
			Melting of II
E II-1	Enantiotropic	II	Melting of II
E II-2	Enantiotropic	II	Solid-state conversion from II to I below T_t
			Solid-state conversion from I to II above T_t
			Melting of II
E II-3	Enantiotropic	II	Solid-state conversion from II to I below T_t
			Melting of I
M I-1	Monotropic	I	Melting of I
M II-1	Monotropic	II	Melting of II
M II-2	Monotropic	II	Solid-state conversion from II to I
			Melting of I
M II-3	Monotropic	II	Melting of II with concurrent recrystallization of I
			Melting of I
M II-4	Monotropic	II	Melting of II
			Recrystallization of I from liquid
			Melting of I

2.5.2 Induced by vapor

Since the vapor pressures of solids are usually very low, this section is more applicable to solvates and hydrates and less relevant to cocrystals. The thermodynamics of solvates clearly demonstrates that they are stable under a subset of conditions where the vapor pressure of the solvent in the environment is higher than some critical values. Under these conditions, the nonsolvated crystal forms are metastable, and driving forces exist for converting them to the solvate forms. On the other hand, under the conditions where the vapor pressure is below this critical value, solvates are metastable, and driving forces exist for converting the solvates to the nonsolvated forms.

Also demonstrated by thermodynamics, activities of the solvents are related to temperature. The same vapor pressure results in lower activities at higher temperatures. Thus, the desolvation phenomenon discussed in Section 2.5.1 is applicable here. However, what is different is the solvation/hydration phenomenon when nonsolvated forms are exposed to a high-solvent vapor pressure. Since exposure only to the organic solvent vapor is quite rare in pharmaceutical development, the discussions here will focus on the role of water vapor and hydrate formation.

RH can vary significantly from location to location and from season to season. Therefore, the phase transition behavior of anhydrates and hydrates has been of particular interest in pharmaceutical development. The kinetic aspect of dehydration has also been studied extensively. From the thermodynamic point of view, when the RH of the environment is closer to RH^c, the driving force for dehydration is lower. Therefore, the dehydration rate should decrease. This is, in fact, observed experimentally for the dehydration of amoxicillin trihydrate.[96] Isothermal dehydration rates were measured at a range of water vapor pressures at 68°C. Dehydration resulted in a poorly crystalline phase, and the phase boundary–controlled kinetic model, under all conditions studied, best described the rates. When the rate constants are plotted against water vapor pressure, regardless of the size of the samples employed, straight lines result, and both extrapolate to a common intercept (within experimental error) on the *x*-axis (Fig. 2.10). This intercept corresponds to the RH^c.

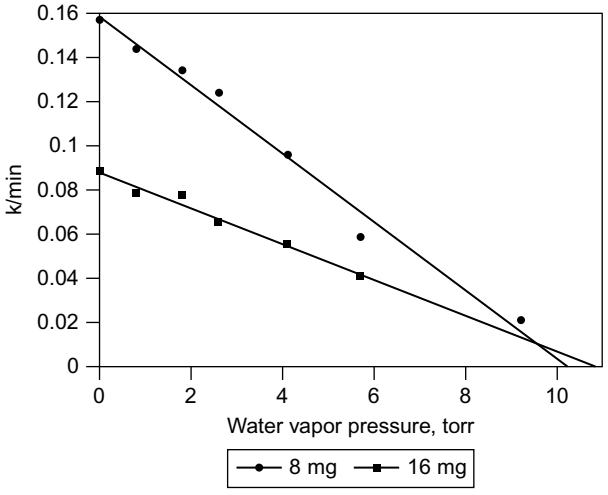


FIGURE 2.10 Plot of the dehydration rate constant of amoxicillin trihydrate as a function of water vapor pressure.[96]

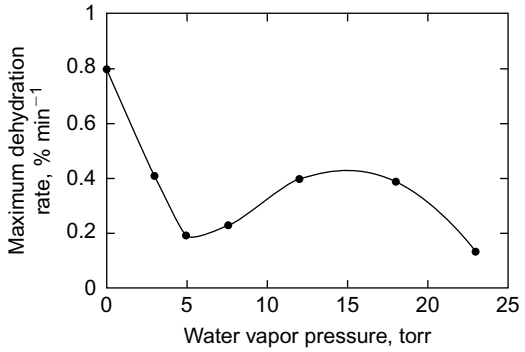


FIGURE 2.11 Maximum dehydration rate of carbamazepine dihydrate as a function of water vapor pressure at 44°C.

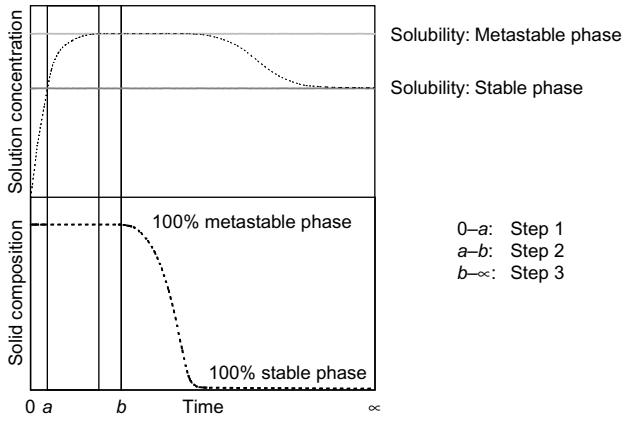


FIGURE 2.12 Concentrations in the solution and the solid compositions as a function of time during the solution-mediated phase transformation.

In the case of carbamazepine dihydrate, however, the rate constant did not follow the same trend (Fig. 2.11).[97] When dehydrating the hydrate at progressively higher water vapor pressure, the rates decrease at first, pass through a minimum, increase and pass through a maximum, and finally decrease again. This phenomenon is known as the *Smith-Topley effect*.[98–100] It has been observed in the dehydration of many other hydrates.[101–107] The cause of this "unusual" behavior is the existence of different dehydration products at low and high water vapor pressures. At extremely low water vapor pressure, the dehydrated phase is amorphous. The water of dehydration will have to diffuse through the amorphous matrix outside each hydrate particle. However, slight increases in water vapor pressure result in the amorphous dehydrated phase undergoing transition to a crystalline phase, therefore creating channels for water to escape and facilitating dehydration. At even higher water vapor pressure, the driving force for dehydration starts to diminish, as in the case of amoxicillin trihydrate.

Exposing anhydrous solids to moisture to form hydrates usually requires excessively higher RH than the RH^c. The probable explanation is that the mechanism of transformation here is not solid-solid, but is instead a solution-mediated transformation. Therefore, bulk solution needs to be present for this mechanism to operate, which is only possible at very high RH.

2.5.3 Induced by solvents

When metastable solids are exposed to solvent (ie, forming a suspension), conversion to the stable phase is facilitated due to a phenomenon called *SMPT*. The presence of the solvent does not change the thermodynamics and stability relationship unless a solvate/hydrate forms with the solvent. However, owing to the much higher mobility in the solution state than in the solid, transformation to the stable phase occurs much faster. This process is analogous to the effect of catalysts for chemical reactions. Mechanistically, SMPT consists of three consecutive steps:[108–110]

1. Initial dissolution of the metastable phase into the solution to reach and exceed the solubility of the stable phase
2. Nucleation of the stable phase
3. Crystal growth of the stable phase coupled with the continuous dissolution of the metastable phase

The schematic representation of concentrations in the solution, as well as the solid compositions as a function of time, are shown in Fig. 2.12 for a typical SMPT process. Kinetically, each step is influenced by a number of factors. Some of the important ones are summarized in Table 2.4. In the overall SMPT process, Step 2 is usually the slowest, and therefore it is the rate-determining step because it involves nucleation. For some fast-nucleating systems, the overall process could be dominated by Step 3. In this case, further differentiation can be made as to whether the process is dissolution or growth controlled.[108] However, differentiation was made based only on rates. Therefore, for the same system (the same solid plus solvent), the controlling step could change by varying some of the experimental conditions (such as temperature, solid/solvent ratio, agitation, particle size, and distribution) and at different extents of transformation.

Many types of phase transitions have been observed for SMPT, including the following:

1. Amorphous to crystalline transition
2. Polymorphic transition
3. Interconversions among the parent and the adducts (ionic, molecular, ionic/molecular)
4. Interconversions between the adducts.

TABLE 2.4 Factors Influencing the Kinetics of the Elementary Steps of SMPT

SMPT step	Factors
1	Temperature (which in turn affects the solubility, diffusion coefficient, and viscosity of the medium)
	Agitation (which in turn affects the thickness of the diffusion layer)
	Initial solution concentration
	Initial particle size and distribution (which in turn affects the initial surface area)
	Solid to solvent ratio
	Solvent medium (which in turn affects the solubility and diffusion coefficient)
	Impurities or additives in the solution (eg, polymers, which in turn affects the diffusion coefficient).
2	Temperature
	Solubility ratio between the two phases
	Solubility
	Nature of the solvent
	Agitation
	Impurities or additives
	Nature and quantity of the contact surfaces (for heterogeneous nucleation)
3	Dissolution component:
	Same as in Step 1, except that the concentration in the solution is higher than the solubility of the table phase
	Crystal growth component
	Temperature
	Solubility ratio between the two phases
	Solubility
	Nature of the solvent
	Agitation
	Available sites for growth
	Number of nuclei
	Presence of secondary nucleation
	Surface roughness
	Impurities or additives in the solution

The occurrences and utility of SMPT span a wide range of pharmaceutically relevant processes including solubility experiments, dissolution (in vitro and in vivo), API crystallization, polymorph/cocrystal screening and characterization, and pharmaceutical processing (such as wet granulation).

2.5.4 Induced by mechanical stresses

When mechanical stresses are applied to solids, they generate many defects. These imperfections create the necessary space and energy to initiate solid-state transitions, such as polymorphic conversion[110–112] and cocrystal formation from a mixture of the corresponding solids.[113,114] Extensive milling, however, often results in a loss of crystallinity and conversion to amorphous material.

2.6 METHODS OF GENERATING SOLIDS

This section briefly summarizes the various methods to generate solids. All crystalline phases are generated by crystallization, where crystallization can occur from gas, liquid (neat or solution), and solid states. Amorphous solids, on the other hand, are generated by the lack of crystallization processes. Hydrates and solvates are generated when crystallized from an environment with high and appropriate water/solvent activities. Ultimately, the driving force for crystallization is the difference in activity (chemical potential), which can be represented differently in gas (vapor pressure), liquid (concentration), and solid states (free energy). Various methods are discussed in some detail next, and they are arranged according to the states of the media through which crystallization occurs: gas, liquid, or solid state.

2.6.1 Through gas

All liquids and solids have vapor pressures at a given temperature. The vapor pressure is a function of temperature, as described by the Clapeyron–Clausius equation. At higher temperatures, the vapor pressure is higher as well. Therefore, by heating a liquid or solid, one can achieve a certain vapor pressure (provided that the substance does not decompose). When this vapor is in contact with a cooler surface, the material in the vapor phase will condense onto the cooler surface, thus generating crystalline polymorphs if nucleation occurs. If nucleation does not occur, the material will condense as an amorphous phase (this is usually termed *vapor deposition*).

Heating is not the only method to generate supersaturation through vapor. It has been demonstrated that reactions in the vapor phase could effectively generate supersaturation and crystallize adducts. One example is the formation of an ephedrine free-base

racemic compound from two opposite enantiomers that are physically separated.[115]

This mode of crystallization is not usually used for hydrates and solvates.

2.6.2 Through liquid

2.6.2.1 Through neat liquid

As shown in Section 2.5.1 earlier in this chapter, polymorphs can be generated via crystallization from the neat liquid at high temperatures. They can also be generated at lower temperatures via crystallization from amorphous phases. The resulting polymorph is a result of the dynamic interplay between the nucleation and growth kinetics of the different polymorphs.

When exposed to moisture or organic solvent vapor, an amorphous phase will absorb solvent molecules from the vapor phase, leading to a reduced glass transition temperature, increased molecular mobility, and ultimately crystallization. This is further discussed in Section 2.7. Potentially hydrates and solvates can crystallize by this method.

2.6.2.2 Through solution

2.6.2.2.1 Solvent evaporation

When evaporating solvent from a solution, the capacity that the solution can accommodate solute is reduced, so supersaturation is eventually created. Upon nucleation, crystalline phases are generated. This evaporative crystallization can also be performed in a solution containing a mixture of solvents. The more volatile solvents, ideally, are good solvents for the solute in this case. When activities of solvent or water are high during the crystallization process, the respective solvates or hydrates may be prepared. On the other hand, amorphous phases are generated if nucleation is sluggish or impaired.

Spray drying is on the extreme side of the rate of solvent removal. Usually, this process generates amorphous material because no time is given for nucleation to occur. However, crystalline phases have been generated via spray drying for some fast-nucleating systems.

2.6.2.2.2 Antisolvent addition

When an antisolvent (which does not provide solvency for the solute) is added to a solution, the composition of the solvent system changes. As a result, the solubility of the solute consequently changes. The supersaturation necessary for nucleation can be generated in this fashion. Again, nucleation leads to crystalline solids, and the lack of nucleation leads to the precipitation of amorphous materials.

The vapor diffusion method involves placing a solution over the vapor of a volatile antisolvent. This method is, in fact, a combination of solvent evaporation and antisolvent addition methods. During the course of the process, the good solvent will evaporate from the solution. Meanwhile, the volatile antisolvent migrates to the solution via the vapor phase.

2.6.2.2.3 Reactive solvent addition

For ionizable compounds, pH adjustment in aqueous media is especially effective at creating supersaturation because of the drastic differences in the solubility. The net effect is to increase the concentration of the less soluble unionized form. Again, nucleation leads to crystalline solids, and the lack of nucleation leads to the precipitation of amorphous materials.

2.6.2.2.4 Temperature gradient

Most organic compounds have temperature-dependent solubility; that is, solubility is usually higher at higher temperatures. Therefore, cooling a saturated solution from higher temperatures will create supersaturation, generating crystalline phases upon nucleation.

Lyophilization is a common method of generating amorphous phases. It is a combination of temperature gradient and solvent evaporation methods. It usually begins with fast cooling of the solution to avoid nucleation. Crystallization of the solvents and subsequent sublimation of solvent crystals effectively remove the solvent from the system, leaving behind an amorphous solid.

2.6.2.2.5 Suspension method

As discussed in Section 2.5.3, the suspension method is particularly effective in generating more stable phases, polymorphs, solvates, or hydrates. Recently, this method has also been extended to screen for cocrystals[44,53,54] and has proven to be successful in screening for novel cocrystals.[116,117]

2.6.3 Through solid

As shown in Section 2.4, solid-state phase transitions are one of the main methods used to generate various solid forms. Heating and cooling, varying RH, and mechanical treatments often result in a change of solid form, thus making solid-state phase transition an attractive means of generating new solids.

2.7 AMORPHOUS DRUGS AND SOLID DISPERSIONS

As mentioned in Sections 2.3 and 2.4 earlier in this chapter, an amorphous phase has higher free energy,

enthalpy, and entropy than the crystalline counterpart, and thus finds application in improving oral bioavailability for BCS class II or IV compounds. However, its unique properties and metastable nature present significant challenges in commercial applications of this concept. In reality, there are only a handful of pharmaceutical products containing amorphous APIs that have been successfully marketed, despite several decades of efforts in research and development. As pointed out by Serajuddin,[118] the limited commercial success reflected challenges from manufacturing difficulties to stability problems. This section aims to briefly survey the fundamental properties of amorphous pharmaceuticals and their solid dispersions.

2.7.1 Characteristics of amorphous phases

2.7.1.1 *Origin of the glass transition*

A number of theories have been proposed to explain the origin of the glass transition and various properties of amorphous materials. The free volume theory and thermodynamic theory are among the most widely accepted.

Free volume theory[119–124] assumes that the volume of a liquid consists of the volume occupied by the constituent molecules and the free volume in which molecules are free to move. The latter exists as voids between molecules and provides room for molecular diffusion. A void can open up due to thermal fluctuations, thus allowing a molecule to move through. The glass transition occurs when the free volume of the system falls below a critical value, and molecular motion by diffusion ceases compared to the experimental timescale.

Dolittle[125] proposed the following relationship between viscosity and free volume:

$$\eta = A\exp(bv_0/v_f) \tag{2.26}$$

Here, v_0 is the molecular volume at 0 K, v_f is the free volume, and A and b are constants. Cohen and Turnbull[122] further related the molecular transport in glasses based on the free volume concept:

$$D = (1/3)(3k_BT/m)^{\frac{1}{2}}\exp(-\gamma v^*/v_f) \tag{2.27}$$

where:

D is the diffusivity,
k_B is the Boltzmann constant,
m is the molecular mass,
v^* is the critical free volume for molecular diffusion.

The free volume theory provides a simple and plausible explanation and is consistent with many experimental observations.

As mentioned previously, the glass transition exhibits the features of a second-order phase transition. Although the glass transition temperature does show kinetic characteristics, it has been suggested that a true second-order transition is masked. The Adam-Gibbs (AG)[126,127] model appears to be able to put together a rigorous explanation of the nature of glass transitions and identify the true second-order transition temperature with the Kauzmann temperature, thus avoiding the entropy crisis. It further relates the molecular relaxation time to the excess entropy of the amorphous phase or the configurational entropy, S_c:

$$\tau = \tau_0\exp\left(\frac{C}{TS_c}\right) \tag{2.28}$$

Here, τ_0 is the relaxation time of molecules at an infinitely high temperature, which conceptually is the relaxation time of the unrestricted molecular relaxation.[128] The overall constant, C, is considered to be a material-dependent parameter that determines the magnitude of the temperature dependence of the relaxation time.

AG theory has been used extensively to explain the properties of the glassy state, and it is more successful than the free volume theory in accounting for the pressure dependence of the glass transition temperature.[129] However, it is not without criticism.[130,131]

2.7.1.2 *Configurational thermodynamic quantities*

The higher heat capacity, entropy, enthalpy, free energy, and other thermodynamic quantities of the amorphous state are due to additional modes of molecular motion in this state, including rotation and translation, which are collectively called *configurations*. Therefore, the excess thermodynamic functions are often called *configurational quantities*. It should be noted that anharmonic vibration might make a significant contribution to those quantities, as noted by Goldstein.[132]

The configurational heat capacity, ΔC_p, entropy, ΔS, enthalpy, ΔH, and free energy, ΔG, can be calculated based on standard thermodynamic treatments:

$$\Delta C_p(T) = C_p^a(T) - C_p^x(T) \tag{2.29}$$

$$\Delta H(T) = H^a(T) - H^x(T) = \Delta H_m + \int_{T_m}^{T}\Delta C_p\,dT \tag{2.30}$$

$$\Delta S(T) = S^a(T) - S^x(T) = \Delta S_m + \int_{T_m}^{T}\frac{\Delta C_p}{T}dT \tag{2.31}$$

$$\Delta G(T) = G^a(T) - G^x(T) = \Delta H(T) - T\Delta S(T) \tag{2.32}$$

Here, the superscripts a, x, and m denote amorphous, crystalline, and melting, respectively.

2.7.1.3 Molecular relaxation in the amorphous state

A glass transition occurs when the timescale of molecular relaxation in the system crosses over the timescale of the observer. Therefore, the glass transition can be studied by changing the timescale of the system under study or the experimental probe. In that sense, the glass transition is purely a kinetic phenomenon, and the molecular relaxation time at the glass transition will apparently depend on the experimental conditions. For example, a fast heating or cooling rate corresponds to a shorter timescale of the probe, which causes the glass transition temperature to shift to a higher temperature (shorter timescale). Periodic perturbations can also be applied, such as in the case of dynamic mechanical analysis, temperature-modulated DSC, dielectric analysis, and rheometric analysis, where the frequency can be changed. Sometimes the comparison among different techniques is not straightforward. However, the average molecular relaxation time at T_g (onset) is approximately 100 seconds for typical DSC scans at 10°C/min.[133]

The temperature dependence of the molecular mobility in the metastable equilibrium supercooled liquid state is usually described by the Vogel-Tammann-Fulcher (VTF) equation,[134–136] as shown here, which also can be approximated by the AG treatment:

$$\tau = \tau_0 \exp\left(\frac{DT_0}{T - T_0}\right) \tag{2.33}$$

where T_0 is the temperature where molecular motion ceases (T_K), while D is a parameter signifying the temperature dependence of the relaxation time constant, or strength of the glass-former. Angell classified glass-formers as either "strong" or "fragile."[137,138] A strong glass has less (lower activation energy) temperature dependence on molecular mobility near T_g than a fragile glass. Molecular mobility in strong glasses typically exhibits Arrhenius-like temperature dependence, while the Arrhenius plot for a fragile glass is significantly nonlinear. A stronger glass is also indicated by smaller change in heat capacity at T_g than a fragile glass. The strength parameter D is higher for strong glasses (such as >100 for SiO_2). Most organic glass-formers are fragile, generally with $D < 15$. Another useful parameter is the fragility parameter, m, which is related to D as follows[139]:

$$m = \frac{\Delta H}{2.303RT_g} = \frac{DT_0}{2.303T_g(1 - T_0/T_g)^2} \tag{2.34}$$

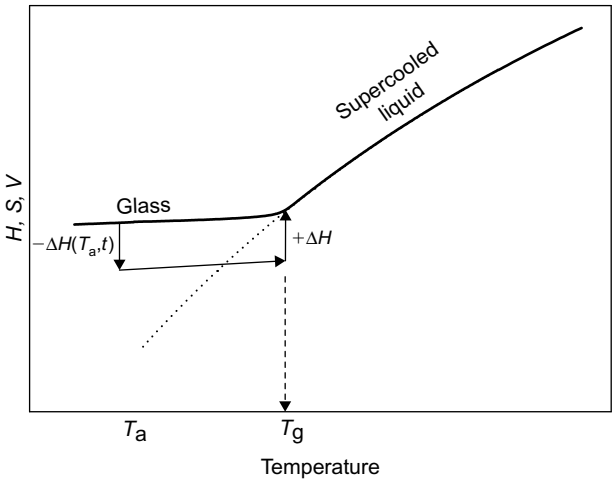


FIGURE 2.13 Schematic representation of the structural relaxation when a glass is annealed at temperature T_a for duration t.

ΔH is the apparent activation energy near T_g, which can be obtained through the heating rate dependence of the glass transition temperature[140,141]:

$$\Delta H = -R\frac{d\ln q}{d(1/T_g)} \tag{2.35}$$

To an experimental observer, molecular motion appears to stop at $T < T_g$ because it takes longer than the timescale of the probe. However, molecular motion can be revealed at a longer timescale. An experimental glass is in a nonequilibrium state. When a glass is stored at temperatures below its T_g, this nonequilibrium state, coupled with mobility over a longer timescale, leads to the so-called physical aging or annealing phenomena. This is due to structural relaxation of the nonequilibrium glass toward the metastable equilibrium or supercooled liquid state (ie, the ideal glass). A schematic representation of this structural relaxation is shown in Fig. 2.13.

According to the AG theory on molecular relaxation in glass-forming liquids, these thermodynamic and dynamic properties are coupled. Therefore, as the excess enthalpy and entropy of the nonequilibrium glass decrease during annealing, the molecular mobility should also decrease. The rate and extent of this decrease in molecular mobility may be important to the physical stability of pharmaceutical glasses.[142]

Molecular relaxation time constant in the annealed glasses can be calculated using the AGV equation[133,143] via the concept of fictive temperature,[133,143–148] T_f:

$$\tau(T, T_f) = \tau_0 \exp\left(\frac{DT_0}{T - (T/T_f)T_0}\right) \tag{2.36}$$

One feature of the structural relaxations in glasses is their nonlinearity or nonexponentiality. Briefly, this

refers to the nonconstant relaxation time constant, which is due to the coupling between the relaxation time and configurational entropy or other quantities. Molecular mobility continuously decreases during the storage of pharmaceutical glasses.[142]

The nonlinearity produces so-called memory effects; therefore, history is an important factor in studying the structural relaxation in glasses. Nonlinearity has been described using the stretched exponential KWW equation:

$$\phi(t) = \exp\left[-\left(\int_0^t \frac{dt'}{\tau(t')}\right)^{\beta}\right] \tag{2.37}$$

Here, $\phi(t)$ is the fraction remaining at time t and β is an empirical parameter indicating the distribution of relaxation time constants ($0 \leq \beta \leq 1$).

2.7.2 Characteristics of amorphous solid dispersions

The term *amorphous solid dispersion* refers to an amorphous drug dispersed in a carrier or matrix in the solid state. Strictly speaking, the term *dispersion* itself implies phase immiscibility. However, in pharmaceutical literature, amorphous solid dispersions are loosely defined, and they also include systems where the drug is dispersed in the matrix at the molecular level (ie, solid solutions). The carriers usually consist of polymers and other "inert" excipients such as surfactants. These carriers could be either amorphous or crystalline. However, amorphous carriers with complete phase miscibility in amorphous drugs offer the most advantages for pharmaceutical applications.

The presence of other components (eg, polymeric carriers) will certainly modify the characteristics of the amorphous drug. Conceptually, two different scenarios could arise: amorphous drug-crystalline polymers, and amorphous drug−amorphous polymers. In the amorphous drug-crystalline polymer system, due to complete phase separation, most likely the properties of the system will be dominated by individual components, except that the interfacial interactions may change certain characteristics of the system (eg, dissolution, aggregation, crystallization, etc.). The following discussion focuses on the other, more common, scenario: amorphous drug−amorphous carrier solid dispersions.

2.7.2.1 Thermodynamic analyses and phase miscibility

Phase miscibility in amorphous solid dispersions is dictated by the thermodynamics of mixing. A decrease in free energy is a necessary condition to initiate mixing. However, it alone does not guarantee phase miscibility.

2.7.2.1.1 Entropy of mixing

Similar to the mixing of small molecules, the mixing of a polymer with a small molecule leads to an increase in entropy, and therefore mixing is favored. However, unlike small molecular mixing, the position availability is limited for polymers due to restraint by the polymer chain, resulting in a smaller ΔS_{mix}. The entropy of mixing for a drug-polymer system can be treated by the Flory−Huggins lattice model:

$$\Delta S_{mix} = R(N_d \ln\phi_d + N_p \ln\phi_p) \tag{2.38}$$

The subscripts d and p designate drug and polymer, respectively; N is the number of sites occupied by the drug or polymer; and ϕ is the volume fraction of drug or polymer in the mixture. In the lattice model treatment, the lattice site is usually defined by the volume of the small molecule. Each polymer molecule then occupies m such sites, where m is the ratio of the molecular volume of the polymer to that of the drug.

2.7.2.1.2 Enthalpy of mixing

Similar to the treatment of regular solution theory, molecular interactions can be built into the lattice model. However, the original regular solution theory merely treated van der Waals interactions, which only accounted for positive deviations from Raoult's law. In the lattice model, specific interactions such as hydrogen bonding were also considered, which play an important role in dictating phase miscibility, and possibly physical stability, in amorphous solid dispersions. The enthalpy of mixing is often expressed as

$$\Delta H_{mix} = RT(N_0 + mN_p)\chi\Phi_0\Phi_p \tag{2.39}$$

where χ is called the Flory−Huggins parameter and χkT represents the enthalpy change of introducing one solvent molecule to the polymer. Depending on the nature of polymer-drug interaction, the enthalpy of mixing can be positive (eg, in the case of only van der Waals interactions) or negative (eg, in the case of hydrogen bonding). Obviously, positive χ causes an increase in enthalpy and free energy, which disfavors mixing.

2.7.2.1.3 Free energy of mixing

Based on the Flory−Huggins model, the total free energy of mixing can be expressed as

$$\begin{aligned}\Delta G_{mix} = RT\,[&N_0 \ln\Phi_0 + N_p \ln\Phi_p \\ &+ (N_0 + mN_p)\chi\,\Phi_0\Phi_p]\end{aligned} \tag{2.40}$$

The entropy of mixing is the only term that always favors mixing. However, this term is much smaller

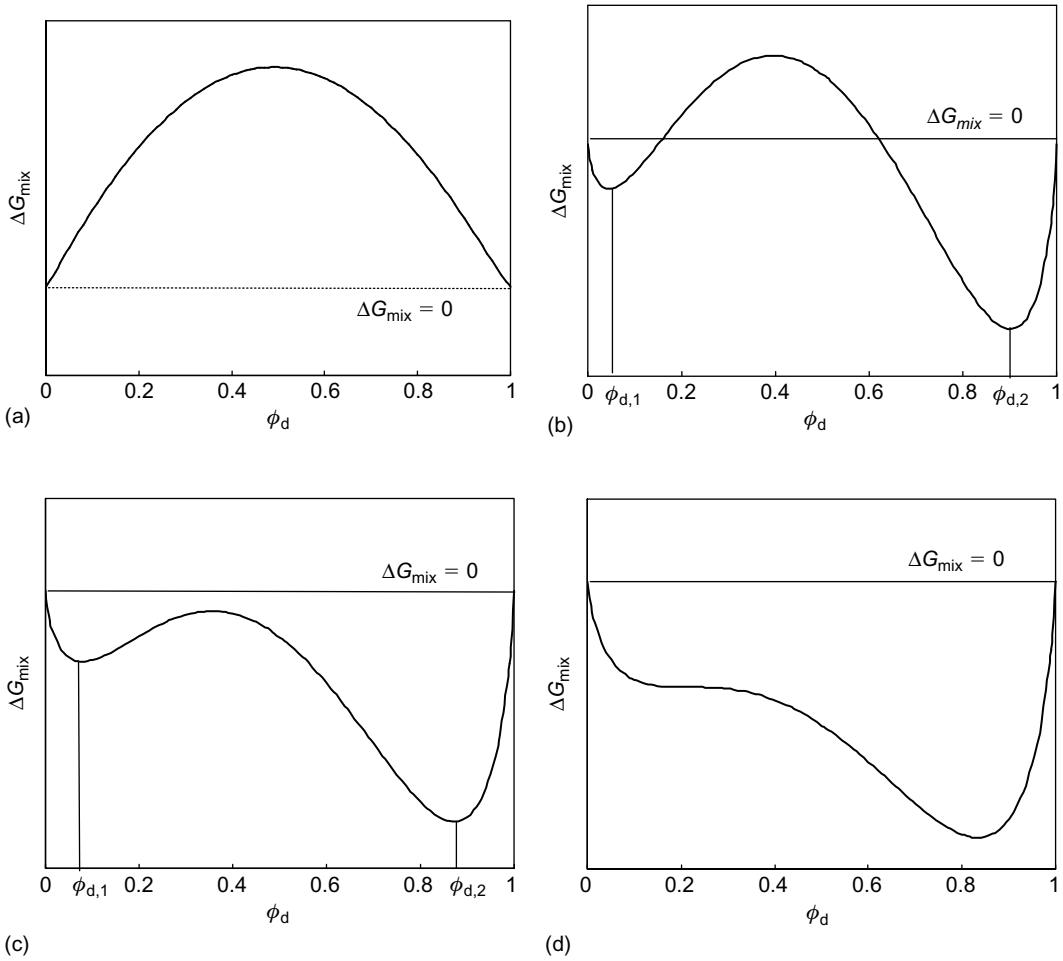


FIGURE 2.14 Schematic free energy of mixing (a) no mixing; (b) partial miscibility; (c) mixing but with phase separation at some compositions; and (d) complete miscibility.

than observed in the mixing of small molecules, and it cannot offset a large positive enthalpy of mixing. Therefore, the free energy of mixing could be negative or positive depending on the sign and magnitude of the Flory–Huggins parameter. If the Flory–Huggins parameter is too large ($\chi > \chi_{crit}$), the free energy of mixing will be positive and no mixing is expected. Specific interactions such as hydrogen bonding make χ negative, so phase miscibility may be favored.

By itself, $\Delta G_{mix} < 0$ does not necessarily imply a single phase. Fig. 2.14 illustrates typical scenarios when mixing small drug molecules with amorphous polymeric carriers. In Fig. 2.14a, $\Delta G_{mix} > 0$ for virtually all compositions; therefore, mixing will not occur and the amorphous drug and polymer will stay in their own pure phases. (Strictly speaking, there will always be a certain extent of mixing. However, in reality, the equilibrium phases can be considered the same as each individual component.) Fig. 2.14b represents partial miscibility. Phase miscibility occurs when $\phi_d < \phi_{d,1}$ or $\phi_d > \phi_{d,2}$. When $\phi_{d,1} < \phi_d < \phi_{d,2}$, the solid dispersion will be phase separated, but not in their individual components, but in a drug-rich phase ($\phi_{d,2}$) and a polymer-rich phase ($\phi_{d,1}$). In Fig. 2.14c, $\Delta G_{mix} < 0$ for all compositions. However, any composition in the range $\phi_{d,1} < \phi_d < \phi_{d,2}$ will still decompose to two local minima represented by $\phi_{d,1}$ and $\phi_{d,2}$, similar to Fig. 2.14b. Therefore, $\Delta G_{mix} < 0$ is only a necessary condition for complete phase miscibility. Fig. 2.14d represents the complete phase miscibility.

2.7.2.2 Molecular mobility in amorphous solid dispersions

When a drug is molecularly dispersed in a polymer, its molecules experience an environment with both other drug molecules and the polymer molecules. The same is true for the polymer. Molecular relaxation in such a dispersed matrix, therefore, is expected to reflect such an environment. The average mobility is usually intermediate between those of the pure components. Assuming additivity of free volumes of each component (ideal mixing), the T_g of an ideal binary

mixture can be estimated using the Gordon–Taylor equation:[149,150]

$$T_{g,mix} = \frac{w_1 T_{g1} + K w_2 T_{g2}}{w_1 + K w_2} \tag{2.41}$$

where w_1 and w_2 are the weight fractions of the components, and K is a constant that can be approximated by the ratio between the respective changes in heat capacity at T_g:

$$K = \Delta C_{p_2} / \Delta C_{p_1} \tag{2.42}$$

To the drug, the presence of the polymer usually causes an increase in T_g, which is often termed an *antiplasticization effect*. In addition, the presence of specific drug-polymer interactions can lead to an increase in T_g above that predicted. However, if the drug-polymer (adhesion) interaction is less than those in drug-drug or polymer-polymer (cohesion) pairs, then T_g is expected to be lower than predicted.

The change in molecular mobility in the amorphous phase by incorporating a second component is of practical relevance to pharmaceuticals other than amorphous solid dispersions. Moisture sorption is vastly increased in an amorphous phase compared to the corresponding crystalline phase. This can be explained by two reasons. The first is that an amorphous phase is more energetic, thus increasing affinity between water and the drug molecules. The second reason is that by removing the crystal lattice, the previous barrier preventing water molecules from entering the crystal lattice vanishes, and water molecules are now able to access almost every host molecule. The latter reason is probably more important. Instead of moisture adsorption on the crystal surface, water molecules form a molecular dispersion with the drug, and the phenomenon is more appropriately called *sorption*. As predicted by the Gordon–Taylor equation, the consequence of moisture sorption into an amorphous phase is a drastically reduced T_g, and increased molecular mobility even for a small amount of moisture sorbed, due to the great plasticization effect of water, which has a T_g of −137°C.

2.7.2.3 Solubility in polymeric matrix

It is apparent that a solid dispersion is metastable only when the drug loading exceeds the solubility limit. Therefore, the solubility of a crystalline drug in a polymeric matrix is of both theoretical and practical importance.

It has been challenging to determine drug solubility in amorphous solid dispersions experimentally, due to the highly viscous nature and slow kinetics of such systems. Theoretically, the solubility could be estimated once the Flory–Huggins interaction parameter becomes known. Methods to derive this parameter in polymer solutions include vapor pressure reduction, osmotic pressure, and other solution techniques, as well as melting-point depression. Most of these solution techniques are difficult to apply to drug-polymer systems. However, the melting-point depression[151,152] can be well adapted to amorphous solid dispersions.[153] This approach is briefly described next.

When a crystalline drug is in equilibrium with a polymer, the chemical potential of the crystalline drug equals that of the dissolved drug in the drug-polymer system. The chemical potential of the drug in solution can be derived based on the free energy of mixing:

$$\begin{aligned}\Delta\mu_d = \mu_d - \mu_d^0 &= \left(\frac{\partial \Delta G_{mix}}{\partial N_d}\right)_{N_p,T,P} \\ &= RT\left[\ln \Phi_d + \Phi_p\left(1 - \frac{1}{m}\right) + \chi \Phi_p^2\right]\end{aligned} \tag{2.43}$$

Here, μ_d^0 is the chemical potential of the pure amorphous drug, which is the same as the molar free energy of the pure amorphous drug at the temperature of the reduced melting point, T_m.

The molar free energy of pure amorphous drug can be related to that of the crystalline drug:

$$\begin{aligned}\mu_d^0(T_m) = G^a(T_m) = G^x(T_m) + \Delta H_m(1 - T_m/T_m^0) \\ + \int_{T_m^0}^{T_m} \Delta C_p\, dT - T\int_{T_m^0}^{T_m} \frac{\Delta C_p}{T}\, dT\end{aligned} \tag{2.44}$$

Here, T_m^0 is the melting temperature of the pure crystalline drug, and $G^x(T_m)$ is the molar free energy of the crystalline drug at T_m, which is identical to its chemical potential, $\mu^x(T_m)$. The following relationship can then be obtained by equating the chemical potential of the drug in polymer and that of the crystalline drug:

$$\begin{aligned}-\frac{\Delta H_m}{R}\left(\frac{1}{T_m} - \frac{1}{T_m^0}\right) - \frac{1}{RT}\int_{T_m}^{T} \Delta C_p\, dT + \frac{1}{R} \\ \times \int_{T_m}^{T} \frac{\Delta C_p}{T}\, dT + = \ln \Phi_0 + \left(1 - \frac{1}{m}\right)\Phi_p + \chi \Phi_p^2\end{aligned} \tag{2.45}$$

The χ-parameter can then be estimated, based on Eq. (2.45) using melting-point depression data. A simplified equation has been used in the literature by ignoring the contribution from excess heat capacity:

$$-\frac{\Delta H_m}{R}\left(\frac{1}{T_m} - \frac{1}{T_m^0}\right) = \ln \Phi_0 + \left(1 - \frac{1}{m}\right)\Phi_p + \chi \Phi_p^2 \tag{2.46}$$

Once the χ parameter becomes known, one can estimate the solubility at lower temperatures.

However, two points need to be made. First, Eq. (2.46) only applies to systems where the amorphous drug and amorphous polymer are in a single phase. As discussed in the previous section, the free energy of mixing will be further reduced if phase separation exists while $\Delta G_{mix}<0$. Second, the solubility obtained from melting-point depression corresponds to that in the rubbery state. However, amorphous solid dispersions are intended to be stored in the glassy state. Therefore, a large discrepancy might indeed be expected.

Solubilization certainly further reduces the driving force for crystallization of the amorphous drug. However, drug solubility in polymers at room temperature is generally quite small; therefore, the tangible impact at normal storage temperatures is possibly limited.

2.7.3 Crystallization of amorphous drugs and dispersions

Crystallization or physical instability is probably one of the greatest challenges for developing amorphous solid dispersions. Although there has been extensive research in this area, consensus has not been reached. A brief summary on the current thinking on this subject follows.

2.7.3.1 Molecular mobility

Molecular mobility plays a key role in physical and chemical stability in amorphous phases, and its importance can be readily understood. For any physical or chemical process, the molecules involved must take appropriate positions and orientations, so that it can proceed.[154] Therefore, molecular motions, such as translation and rotation, are essential in any physical or chemical change of the system. In solution or in the melt, molecular translation and rotation are usually faster than the physical or chemical change itself, so that the overall rate is independent of diffusion of the reactant or product. However, in the solid state, molecular diffusion is orders of magnitudes slower than in solution and, in most cases, can be comparable to, or even slower than, the rate of reaction. It is likely that molecular motion becomes partially or completely rate-limiting in solid-state reactions.

The partial correlation between molecular mobility and physical/chemical stability has been demonstrated for a number of pharmaceutical systems.[155–159] The crystallization rates of amorphous nifedipine in the presence of different excipients were correlated to the spin-lattice relaxation time of D_2O in these matrices.[155] Differences in crystallization rates of amorphous nifedipine, phenobarbital, and flopropione were attributed to their differences in molecular mobility.[156] In a follow-up study, the temperature dependence of physical stability of amorphous nifedipine and phenobarbital were correlated to a large extent to the temperature dependence of molecular relaxation time constant.[157] Physical stability of a lyophilized monoclonal antibody in sucrose and trehalose was correlated to molecular mobility measured by enthalpy relaxation, but not the glass transition temperature.[158] It was also shown that the fragility of the formulations played a role. Chemical degradation of amorphous quinapril hydrochloride demonstrated significant correlation with molecular relaxation time constant at temperatures below T_g.[159] Bimolecular reaction rates of acetyl transfer, as well as the Maillard reaction in lyophilized aspirin-sulfadiazine formulations, were concluded to be predictable based on molecular mobility.[160] While the chemical stability of lyophilized insulin showed correlation with the molecular mobility in trehalose formulation, the contribution of molecular mobility was found to be negligible in a PVP formulation.[161,162]

To maintain adequate physical stability, storage below T_g is apparently needed. However, unlike the metastable equilibrium supercooled state, molecular mobility in the glassy state is much higher and is less dependent on temperature. Therefore, the advantage achieved by just storing an amorphous formulation at temperatures far below T_g is not as great as expected, based on that predicted using the VTF equation. Along the same lines, the suggestion of storage at the Kauzmann temperature (typically $T_g=-50°C$, but varying depending on the fragility) will likely not provide the intended consequences because realistic glasses possess much higher mobility than ideal glasses, and molecular mobility at T_K is still significant. Nevertheless, raising the T_g of amorphous dispersions by incorporating a high T_g excipient is still a sound tactic, as implied by the Gordon–Taylor equation. However, the limitations of this approach need to be kept in mind. T_g is not necessarily a predictor of physical stability.[158] In addition, compromises may be necessary because a higher process temperature will usually be required for higher T_g formulations if the melt method is employed in manufacturing.

The rate of crystallization may also decouple from the viscosity of glasses. Hikima et al.[163] studied the *o*-terphenyl system and observed that crystal growth increased abruptly near T_g and below, orders of magnitude faster than expected from diffusion-controlled growth. Most recently, Yu and coworkers identified a few pharmaceutical systems exhibiting similar behavior, including indomethacin,[164] nifedipine,[165] and the ROY system.[166,167] Many questions remain to be answered; however, this glass–crystal mode of crystal growth appeared to be related to the precursors present in liquid and crystal structure (polymorph).[166,167]

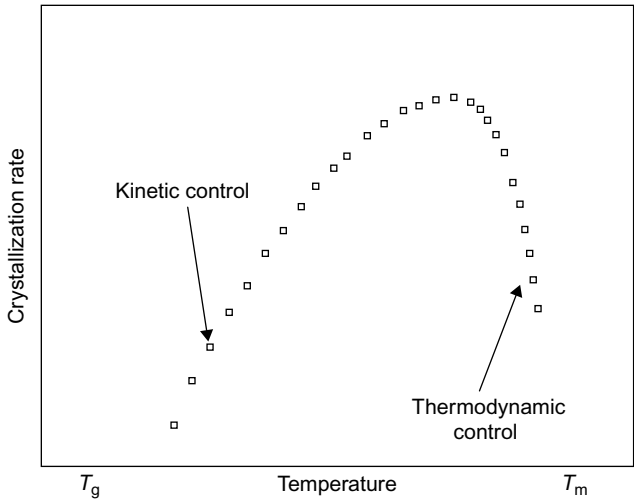


FIGURE 2.15 Schematic representation of crystallization rate from supercooled liquids.

2.7.3.2 *Free energy driving force*

In addition to the molecular mobility, the role of the free energy driving force has been historically discussed when dealing with crystallization from supercooled melts.[168] Mobility determines how fast a molecule can impinge on the surface of a growing crystal. However, not every impingement leads to successful incorporation of the molecule. The probability of the molecule not coming back to the liquid phase is related to the free energy driving force: the higher the free energy difference, the more likely the molecule will stay in the crystal lattice. The relationship can be described as follows:[164]

$$u = (k/\eta)[1-\exp(-\Delta G/RT)] \tag{2.47}$$

where:

η is the viscosity (proportional to molecular relaxation time constant),

ΔG is the free energy driving force,

R is a gas constant,

k is a constant.

While the viscosity of a supercooled liquid can be described by an empirical equation such as VTF, ΔG is approximately proportional to the degree of supercooling, $\Delta T = T_m - T$. Based on this relationship, the temperature for the maximum crystallization rate falls between the glass transition temperature and the melting temperature, as depicted in Fig. 2.15.

2.7.3.3 *Configurational entropy*

Considering the metastable nature of an amorphous phase (ie, $\Delta G = G^a - G^x > 0$), it is interesting that the amorphous phase has both higher enthalpy and higher entropy. From the thermodynamic relationship $\Delta G = \Delta H - T\Delta S$, it can be concluded that the higher free energy of amorphous phase is due to the enthalpy factor, while the entropy term partially offsets the enthalpy term, which actually reduces the overall free energy. It is apparent that the free energy and the enthalpy of amorphous state facilitate reversion to the crystalline form. The entropy term appears to behave somewhat differently.

As discussed previously, crystallization from supercooled melts are generally discussed in the context of molecular mobility and free energy driving force. For crystallization near and below T_g, which is generally the case where physical stability of amorphous pharmaceuticals is concerned, crystallization is mainly dictated by molecular mobility. Indeed, molecular mobility is generally accepted as a principal factor governing the physical stability of amorphous phases.[62,169,170] One simple fact is that crystallization of amorphous phases proceeds much faster in the rubbery state compared to the glassy state. As briefly noted, molecular mobility has been correlated partially or fully to physical/chemical stability in a number of pharmaceutical systems.

Nevertheless, significant differences in crystallization behavior are observed across compounds, and these cannot be explained by mobility alone. For example, some amorphous phases crystallize almost immediately at the glass temperature (eg, acetaminophen[171]), some crystallize below T_g in a relatively short time (eg, griseofulvin[172] and nifedipine[156,157]), while others are quite stable. For some of the more stable amorphous phases, such as ritonavir,[171] crystallization in the glassy state is not observed, and it does not proceed at a significant rate above T_g without seeding. Because T_g corresponds to the temperature where the molecular relaxation time of the amorphous phase is equivalent to the experimental timescale, it is clear that the same molecular mobility results in different crystallization tendencies.

The authors have explored the physical stability of a diverse group of pharmaceutically relevant compounds, attempting to identify thermodynamic quantities contributing to crystallization tendency. Through this analysis, it was concluded that the calorimetric configurational entropy also played an important role.[142,171]

In simple terms, configurational entropy is a measure of the difference in the number of configurations between the amorphous and crystalline phases. For molecules in the amorphous state to crystallize, the molecule has to pack into a defined crystal lattice with a defined configuration or orientation. Therefore, the higher the configurational entropy, the lower the probability that the molecules will assume the desirable orientation for packing into the crystal lattice. Hence, a metastable amorphous compound with large

configurational entropy tends to show greater physical stability. This is analogous to the common experience of chemists and biochemists that large molecules with numerous rotatable bonds are generally difficult to crystallize from solution. It has been hypothesized that configurational entropy serves as a thermodynamic measurement of the probability of nucleation, while the molecular mobility dictates the rate at which a molecule can change its configurations and serves as a kinetic measurement of nucleation. When these two quantities were evaluated with respect to molecular structure, it was possible to provide an assessment of physical stability. Therefore, in relative terms, small, rigid molecules will tend to crystallize more easily than large, flexible compounds.

Recently, Ediger et al.[173] surveyed the crystallization rates of a number of inorganic and organic systems and identified a correlation between rate of crystallization and entropy difference. In light of this new development, it seems appropriate that roles of mobility, entropy, and free energy during crystallization can be discussed in the context of three consecutive events during crystallization: mobility determines the rate of molecule approaching a crystal surface; entropy relates to the probability of that approaching molecule having the desirable configurations; and the free energy dictates the probability of that molecule not returning to the liquid phase. A more or less quantitative relationship may be represented as

$$u = (k/\eta^{\xi}) \exp(-\Delta S/R)[1-\exp(-\Delta G/RT)] \qquad (2.48)$$

Here, ξ is a constant describing the extent of correlation between crystallization rate and diffusivity.

2.7.3.4 *Crystallization inhibition*

Historically, polymers have been found to inhibit crystallization from solution. Among those pharmaceutically relevant ones, PVP has been extensively studied.[174–178] Other crystallization inhibitors include Pluronic F68,[179] cellulose derivatives,[180] and cyclodextrins.[181]

Polymeric carriers or other additives have also been shown to inhibit crystallization from amorphous solid dispersions. Thermodynamics dictates that amorphous drug would crystallize if given sufficient time. However, from the perspective of a formulation scientist, it is of great benefit if one can delay this crystallization process to obtain adequate shelf life. There are many examples of crystallization inhibition in amorphous solid dispersions;[182–188] however the mechanisms of such effects remain to be determined.

When dealing with solid dispersions, a key parameter is the molecular mobility or, approximately, T_g. The antiplasticization effect of polymeric carriers is considered important and could conceptually lead to a more stable amorphous phase than the drug itself. However, this factor alone is not entirely responsible for some experimental observations. For example, amorphous solid dispersions of felodipine with PVP, hydroxypropyl methylcellulose acetate succinate (HPMCAS), and hydroxypropyl methyl cellulose (HPMC)[188] showed similar stability profiles, while only its dispersion with PVP showed a significant increase in T_g for polymer loading up to 25%.

Specific interactions between drug and polymers were also considered to be important. For example, the inhibition of crystallization of amorphous indomethacin by PVP and PVP-*co*-PVA was ascribed to the hydrogen bonding interaction between the polymers and drug molecules.[183] PVP was found to inhibit the crystallization of indomethacin more with larger particles than smaller ones,[184] and was consistent with the observation of predominantly surface nucleation. An ion-dipole interaction was proposed to be partially responsible for the crystallization inhibition in solid dispersions of PVP[185] and a developmental drug, which could not be attributed to mobility alone. The increase in physical stability of amorphous nifedipine in ethylcellulose and Eudragit RL dispersions were also explained by the hydrogen-bonding interaction between drugs and polymers.[186] Strong interactions were observed in acetaminophen/aminoacetanilide and PVP/polyacryric acid polymers, which led to improved physical stability[187] and correlated with the strength of the specific interaction.[189] However, the felodipine dispersions with PVP, HPMC, and HPMCAS showed similar inhibitory effects, although the hydrogen-bonding strengths followed PVP > HPMC > HPMCAS.[188] These findings may not appear to be consistent. The situation is quite complex with the presence of specific interactions. On the one hand, strong drug-polymer interactions lead to greater reductions in molecular mobility than predicted by free volume additivity alone. On the other hand, the presence of certain intermolecular structures between drug and polymer molecules due to specific interactions may create a kinetic barrier for the amorphous drug to crystallize. Additionally, specific interactions decrease the free energy of the system. All these factors appear to work toward stabilization, albeit maybe to different extents. However, it is difficult to delineate the individual contributions.

While it is important to maintain the physical stability of amorphous dispersions during storage, it is also useful to utilize similar concepts for solid dispersions during in vivo dissolution. The higher dissolution rate and apparent solubility of an amorphous drug usually causes supersaturation during in vivo dissolution, and therefore may lead to precipitation in the GI tract and compromise oral

bioavailability. If such a crystallization-inhibitory polymer is incorporated into the amorphous solid dispersion, then the in vivo precipitation may be delayed or completely eliminated, resulting in much improved oral absorption. It is ideal if the polymeric carrier can function as a crystallization inhibitor, both during storage and during in vivo dissolution. However, the mechanisms may not be the same, and different inhibitors may indeed be necessary. A few examples have been reported. Various itraconazole solid dispersions were prepared by hot-melt extrusion and compared with the commercial drug Sporanox in a clinical study.[66] The bioavailability relative to Sporanox was 102.9%, 77.0%, and 68% for the HPMC, Eudragit E 100, and Eudragit E 100-poly(vinylpyrollidone-*co*-vinylacetate) solid dispersions, respectively, in opposite rank order of their in vitro dissolution data. The reason was speculated to be due to the crystallization-inhibitory role of HPMC during dissolution. Similarly, an in vitro dissolution study indicated that supersaturation was maintained for an amorphous solid dispersion of albendazole in hydroxypropyl methylcellulose phthalate (HP-55), but not with HPMC,[190] which lead to its higher in vivo performance in rabbits.[190]

2.8 SPECIAL TOPICS

2.8.1 Polymorph screening and stable form screening

Polymorph screening is a practice focusing on the discovery of crystalline solids, mostly polymorphs. However, it includes the search for hydrates and solvates as well. Historically, polymorph screening usually included the following approaches:

1. Recrystallization under various conditions: the crystallinity of the original solid is removed by dissolving in solutions or melting. Crystallinity is then regenerated under a variety of conditions (such as temperature, solvent, supersaturation, additives, etc.) to see whether it results in other crystal forms.
2. Examining the solid-solid transitions induced by heat, humidity, mechanical stresses, etc.

Although it is valuable to screen and discover all crystal forms for a particular compound, finding the most stable polymorph is practically the most important aspect in polymorph screening. A polymorph-screening procedure that provides a greater probability of finding the stable polymorph was formulated.[191] The use of a suspension/slurry and the principle of SMPT to facilitate this process was described in detail as follows:

1. Recrystallization:
   a. Use the available polymorph.
   b. Ensure complete dissolution of the solid, to eliminate any seeds of the original polymorph.
   c. Recrystallize at a low rate, either at a low cooling rate or at a low rate of solvent evaporation.
   d. Use a wide range of final temperatures. If cooling crystallization is employed, the induction time and the rate of transformation of the metastable phase to the stable phase may vary with temperature.
2. After recrystallization, add a small amount of the original polymorph to the suspension and stir for various times, and examine the solid phase regularly. The length of time depends on the compound of interest, and the longer the better, provided that it has satisfactory chemical stability in the solution. The minimum time recommended is 7 days.
3. A range of solvents, with a variety of different characteristics such as hydrogen bonding, dielectric constant, solvency for the compound, should be employed.

This procedure improves the chance of finding the stable polymorph when a new pharmaceutical candidate is under investigation. Although slightly more time and effort are required, this procedure is technically easy, and the outcome may easily outweigh the effort.

Recently, one variation of this procedure, where initial crystallization was not performed, was reported being used in the routine early screening for stable phases.[192] Encouraging statistical data were reported by the authors:

> Based on limited statistics ($n = 43$), approximately 26% of early discovery compounds (received as crystalline phases) converted to more stable polymorphs during the stable polymorph screen. To date, more stable polymorphs have not been observed for any of those 43 candidates that have advanced into development.

2.8.2 High-Throughput crystallization

Automation in crystallization has been on the rise in recent years and has shown great promise in the areas of polymorph screening and crystal engineering; that is, searching for cocrystals of better physicochemical properties.[193–196] There are several advantages of robotic crystallization over conventional manual crystallization. First, the robot is designed to perform tasks in a parallel fashion; therefore, it is able to

examine the various parameters quickly. Second, because of the rapidity, robotic crystallization is able to explore parameters influencing crystallization more thoroughly, leading to a better understanding. Third, robotic crystallization requires less material. Fourth, because of the thoroughness, robotic crystallization may produce crystal forms that are otherwise not accessible by conventional crystallization.

2.8.3 Miniaturization in crystallization

Another trend in the crystallization of pharmaceutical materials is miniaturization. Reducing the material needed is essential for studies during early development, where the cost of API is extremely high and availability is scarce. Several studies were conducted using capillary tubes.[197–199] In addition to the benefit of reduced material consumption (5–50 μL of solution), the authors reported the discovery of previous unknown polymorphs owing to crystallization in a confined space in some cases.

Most recently, microfluidics technology that is capable of scaling down further by 10- to 100-fold has been applied to crystallization. Although most of the current applications are toward the crystallization of proteins,[200–203] the extension to pharmaceutically relevant small molecules has already begun.[204–206]

References

1. Byrn SR, Pfeiffer RR, Stowell JG. *Solid-state chemistry of drugs*. 2nd ed. West Lafayette, IN: SSCI, Inc; 1999.
2. Bernstein J. *Polymorphism in molecular crystals*. Oxford, UK: Oxford University Press; 2002.
3. Brittain HG. *Polymorphism in pharmaceutical solids*. *Drugs Pharm Sci*, vol. 95. New York, NY: Dekker; 1999.
4. Brittain HG. *Physical characterization of pharmaceutical solids*. *Drugs Pharm Sci*, vol. 70. New York, NY: Dekker; 1995.
5. Hilfiker R. *Polymorphism: in the pharmaceutical industry*. KGaA, Weinheim, Germany: Wiley-VCH Verlag GmbH & Co; 2006.
6. Brittain HG. *Spectroscopy of pharmaceutical solids*. *Drugs Pharm Sci*, vol. 160. New York, NY: Taylor & Francis; 2006.
7. Stahl PH, Wermuth CG. *Handbook of pharmaceutical salts: properties, selection, and use*. Zürich, Switzerland and Wiley-VCH, Weinheim, Germany: VHCA; 2002.
8. Characterization of the solid state. In: Suryanarayanan R, Byrn SR, editors. *Adv Drug Del Rev*, 48. 2001. p. 1–136.
9. Raw AS, Yu LX, editors. Pharmaceutical solid polymorphism in drug development and regulation. *Adv Drug Del Rev* 2004; **56**:235–414.
10. Polymorphism in crystals. In: Rogers RD *Cryst Grow Design*, 4. 2004. p. 1085–441.
11. Polymorphism in crystals. In: Rogers RD *Cryst Grow Design*, 3. 2003. p. 867–1040.
12. Facets of polymorphism in crystals. In: Matzger AJ *Cryst Grow Design*, 8. 2008. p. 2–161.
13. Haleblian JK. Characterization of habits and crystalline modification of solids and their pharmaceutical applications. *J Pharm Sci* 1975;**64**:1269–88.
14. Desiraju GR. Crystal and co-crystal. *CrystEngComm* 2003;**5**:466–7.
15. Dunitz JD. Crystal and co-crystal: a second opinion. *CrystEngComm* 2003;**5**:506.
16. Aakeröy CB, Salmon DJ. Building co-crystals with molecular sense and supramolecular sensibility. *CrystEngComm* 2005;**7**:439–48.
17. Lara-Ochoa F, Espinosa-Perez G. Crystals and patents. *Cryst Grow Design* 2007;**7**:1213–15.
18. Bond AD. What is a co-crystal? *CrystEngComm* 2007;**9**:833–4.
19. Lara-Ochoa F, Espinosa-Perez G. Cocrystals definitions. *Supramol Chem* 2007;**19**:553–7.
20. Schmidt AC, Schwarz I. Solid-state characterization of non-stoichiometric hydrates of ester-type local anaesthetics. *Int J Pharm* 2006;**320**:4–13.
21. Authelin JR. Thermodynamics of non-stoichiometric pharmaceutical hydrates. *Int J Pharm* 2005;**303**:37–53.
22. Stephenson GA, Diseroad BA. Structural relationship and desolvation behavior of cromolyn, cefazolin and fenoprofen sodium hydrates. *Int J Pharm* 2000;**198**:167–77.
23. Chen LR, Young Jr. VG, Lechuga-Ballesteros D, Grant DJW. Solid-state behavior of cromolyn sodium hydrates. *J Pharm Sci* 1999;**88**:1191–200.
24. Reutzel SM, Russell VA. Origins of the unusual hygroscopicity observed in LY297802 tartrate. *J Pharm Sci* 1998;**87**:1568–71.
25. Variankaval N, Wenslow R, Murry J, Hartman R, Helmy R, Kwong E, et al. Preparation and solid-state characterization of nonstoichiometric cocrystals of a phosphodiesterase-IV inhibitor and L-tartaric acid. *Cryst Grow Design* 2006;**6**:690–700.
26. Yu L. Inferring thermodynamic stability relationship of polymorphs from melting data. *J Pharm Sci* 1995;**84**:966–74.
27. Yu L, Huang J, Jones KJ. Measuring free-energy difference between crystal polymorphs through eutectic melting. *J Phys Chem B* 2005;**109**:19915–22.
28. Yu L, Stephenson GA, Mitchell CA, Bunnell CA, Snorek SV, Bowyer JJ, et al. Thermochemistry and conformational polymorphism of a hexamorphic crystal system. *J Am Chem Soc* 2000;**122**:585–91.
29. Chen S, Guzei IA, Yu L. New polymorphs of ROY and new record for coexisting polymorphs of solved structures. *J Am Chem Soc* 2005;**127**:9881–5.
30. Reutzel-Edens SM, Russell VA, Yu L. Molecular basis for the stability relationships between homochiral and racemic crystals of tazofelone: a spectroscopic, crystallographic, and thermodynamic investigation. *Perkin* 2000;**25**:913–24.
31. Noyes AA, Whitney WR. The rate of solution of solid substances in their own solutions. *J Am Chem Soc* 1897;**19**:930–4.
32. Burger A, Ramberger R. On the polymorphism of pharmaceuticals and other molecular crystals. I. Theory of thermodynamic rules. *Mikrochimi Acta [Wien]* 1979;**II**:259–71.
33. Burger A, Ramberger R. On the polymorphism of pharmaceuticals and other molecular crystals. II. Applicability of thermodynamic rules. *Mikrochimi Acta [Wien]* 1979;**II**:273–316.
34. Zhu H, Yuen C, Grant DJW. Influence of water activity in organic solvent 1 water mixtures on the nature of the crystallizing drug phase. 1. *Theophylline Int J Pharm* 1996;**135**:151–60.
35. Zhu H, Grant DJW. Influence of water activity in organic solvent 1 water mixtures on the nature of the crystallizing drug phase. 2. Ampicillin. *Int J Pharm* 1996;**139**:33–43.
36. Khankari R, Chen L, Grant DJW. Physical characterization of nedocromil sodium hydrates. *J Pharm Sci* 1998;**87**:1052–61.
37. Bandyopadhyay R, Erixon K, Young VG, Grant DJW. Effects of water activity on recrystallized L-lysine monohydrochloride. In: *World Congress on Particle Technology 3*, July 6–9. Brighton, UK; 1998.p. 1889–1909.

38. Beckmann W, Winter G. Stability of the hydrate and anhydrate of pyrazzocarnil hydrochloride in water-ethanol mixtures and in moist air. In: *14th International Symposium on Industrial Crystallization*, September 12–16. Cambridge, UK; 1999. p. 236–245.
39. Wadso L, Markova N. Comparison of three methods to find the vapor activity of a hydration step. *Eur J Pharm Biopharm* 2001;**51**:77–81.
40. Ticehurst MD, Storey RA, Watt C. Application of slurry bridging experiments at controlled water activities to predict the solid-state conversion between anhydrous and hydrated forms using theophylline as a model drug. *Int J Pharm* 2002;**247**:1–10.
41. Sacchetti M. Determining the relative physical stability of anhydrous and hydrous crystal forms of GW2016. *Int J Pharm* 2004;**273**:195–202.
42. Han J, Zhang GGZ, Grant DJW, Suryanarayanan R. *Physical stability of pharmaceutical hydrates*, in preparation.
43. Suzuki E, Shimomura K, Sekiguchi K. Thermochemical study of theophylline and its hydrate. *Chem Pharm Bull* 1989;**37**:493–7.
44. Zhang GGZ, Henry RF, Borchardt TB, Lou X. Efficient co-crystal screening using solution-mediated phase transformation. *J Pharm Sci* 2007;**96**:990–5.
45. Kauzmann W. The nature of the glassy state and the behavior of liquids at low temperatures. *Chem Rev* 1948;**43**:219–56.
46. Ehrenfest P. Phase changes in the ordinary and extended sense classified according to the corresponding singularities of the thermodynamic potential. *Proc Acad Sci Amsterdam* 1933;**36**:153–7.
47. The United States Pharmacopeia. *30th revision*. Rockville, MD, USA: United States Pharmacopeial Convention; 2007.
48. Haleblian JK, Mccrone W. Pharmaceutical applications of polymorphism. *J Pharm Sci* 1969;**58**:911–29.
49. Threlfall TL. Analysis of organic polymorphs. A review. *Analyst (Cambridge, United Kingdom)* 1995;**120**:2435–60.
50. Giron D. Thermal analysis and calorimetric methods in the characterization of polymorphs and solvates. *Thermochim Acta* 1995;**248**:1–59.
51. Grant DJW. Theory and origin of polymorphism. In: Brittain HG, editor. *Polymorphism in pharmaceutical solids*. New York, NY: Marcel Dekker; 1999. p. 1–33.
52. Pudipeddi M, Serajuddin ATM. Trends in solubility of polymorphs. *J Pharm Sci* 2005;**94**:929–39.
53. Nehm SJ, Rodríguez-Spong B, Rodríguez-Hornedo N. Phase solubility diagrams of cocrystals are explained by solubility product and solution complexation. *Cryst Grow Design* 2006;**6**:592–600.
54. Rodríguez-Hornedo N, Nehm SJ, Seefeldt KF. Reaction crystallization of pharmaceutical molecular complexes. *Mol Pharm* 2006;**3**:362–7.
55. Hancock BC, Parks M. What is the true solubility advantage for amorphous pharmaceuticals? *Pharm Res* 2000;**17**:397–403.
56. Imaizumi H. Stability and several physical properties of amorphous and crystalline forms of indomethacin. *Chem Pharm Bull* 1980;**28**:2565–9.
57. Miyazaki S, Hori R, Arita T. Physico-chemical and gastrointestinal absorption of some solid phases of tetracycline. *Yakugaku Zasshi* 1975;**95**:629–33.
58. Sato T, Okada A, Sekiguchi K, Tsuda Y. Difference in physicopharmaceutical properties between crystalline and non-crystalline 9,3″-diacetylmidecamycin. *Chem Pharm Bull* 1981;**29**:2675–82.
59. Mullins J, Macek T. Some pharmaceutical properties of novobiocin. *J Am Pharm Ass Sci Ed* 1960;**49**:245–8.
60. Shefter E, Higuchi T. Dissolution behavior of crystalline solvated and nonsolvated forms of some pharmaceuticals. *J Pharm Sci* 1963;**52**:781–91.
61. Stagner WC, Guillory JK. Physical characterization of solid iopanoic acid forms. *J Pharm Sci* 1979;**68**:1005–9.
62. Hancock BC, Zografi G. Characteristics and significance of the amorphous state in pharmaceutical systems. *J Pharm Sci* 1997;**86**:1–12.
63. Law D, Krill SL, Schmitt EA, Fort JJ, Qiu Y, Wang W, et al. Physicochemical considerations in the preparation of amorphous ritonavir-poly(ethylene glycol) 8000 solid dispersions. *J Pharm Sci* 2001;**90**:1015–25.
64. Law D, Schmitt EA, Marsh KC, Everitt EA, Wang W, Fort JJ, et al. Ritonavir-PEG 8000 amorphous solid dispersions: in vitro and in vivo evaluations. *J Pharm Sci* 2004;**93**:563–70.
65. Parke-Davis. Rezulin® [package insert]; 1999.
66. Six K, Daems T, De Hoon J, Van Hecken A, Depre M, Bouche M-P, et al. Clinical study of solid dispersions of itraconazole prepared by hot-stage extrusion. *Eur J Pharm Sci* 2005;**24**:179–86.
67. Yamashita K, Nakate T, Okimoto K, Ohike A, Tokunaga Y, Ibuki R, et al. Establishment of new preparation method for solid dispersion formulation of tacrolimus. *Int J Pharm* 2003;**267**:79–91.
68. Breitenbach J. Melt extrusion can bring new benefits to HIV therapy. *Am J Drug Del* 2006;**4**:61–4.
69. Valeant Pharmaceuticals International. Cesamet® [package insert]; 2006.
70. Zografi G. States of water associated with solids. *Drug Dev Ind Pharm* 1988;**14**:1905–26.
71. Newman AW, Reutzel-Edens SM, Zografi G. Characterization of the "hygroscopic" properties of active pharmaceutical ingredients. *J Pharm Sci* 2007;**97**:1047–59.
72. Ahlneck C, Zografi G. The molecular basis of moisture effects on the physical and chemical stability of drugs in the solid state. *Int J Pharm* 1990;**62**:87–95.
73. Cohen MD, Green BS. Organic chemistry in the solid state. *Chem Br* 1973;**9**:490–7.
74. Schmidt GMJ. Topochemistry. III. The crystal chemistry of some trans-cinnamic acids. *J Chem Soc* 1964;**43**:2014–21.
75. Byrn SR, Sutton PA, Tobias B, Frye J, Main P. Crystal structure, solid-state NMR spectra, and oxygen reactivity of five crystal forms of prednisolone tert-butylacetate. *J Am Chem Soc* 1988;**110**:1609–14.
76. Byrn SR, Kessler DW. The solid state reactivity of the crystal forms of hydrocortisone esters. *Tetrahedron* 1987;**43**:1335–43.
77. Sukenik CN, Bonapace JAP, Mandel NS, Lau P-Y, Wood G, Bergman RG. A kinetic and x-ray diffraction study of the solid state rearrangement of methyl p-dimethylaminobenzenesulfonate. Reaction rate enhancement due to proper orientation in a crystal. *J Am Chem Soc* 1977;**99**:851–8.
78. Sukenik CN, Bonopace JA, Mandel NS, Bergman RC, Lau PY, Wood G. Enhancement of a chemical reaction rate by proper orientation of reacting molecules in the solid state. *J Am Chem Soc* 1975;**97**:5290–1.
79. Leeson LJ, Mattocks AM. Decomposition of aspirin in the solid state. *J Am Pharm Ass* 1958;**47**:329–33.
80. Chen X, Li T, Morris KR, Byrn SR. Crystal packing and chemical reactivity of two polymorphs of flufenamic acid with ammonia. *Mol Cryst Liquid Cryst Sci Technol* 2002;**381**:121–31.
81. Kontny MJ, Grandolfi GP, Zografi G. Water vapor sorption of water-soluble substances: studies of crystalline solids below their critical relative humidities. *Pharm Res* 1987;**4**:104–12.
82. Wardrop J, Law D, Qiu Y, Engh K, Faitsch L, Ling C. Influence of solid phase and formulation processing on stability of Abbott-232 tablet formulations. *J Pharm Sci* 2006;**95**:2380–92.
83. Hess H. Tablets under the microscope. *Pharm Tech* 1978;**2**:38–57.
84. Roberts RJ, Rowe RC. The Young's modulus of pharmaceutical materials. *Int J Pharm* 1987;**37**:15–18.

85. Duncan-Hewitt WC, Weatherly GC. Modeling the uniaxial compaction of pharmaceutical powders using the mechanical properties of single crystals. II: brittle materials. *J Pharm Sci* 1990;**79**:273–8.
86. Danielson DW, Morehead WT, Rippie EC. Unloading and post-compression viscoelastic stress versus strain behavior of pharmaceutical solids. *J Pharm Sci* 1983;**72**:342–5.
87. Martino PD, Guyot-Hermann A-M, Conflant P, Drache M, Guyot J-C. A new pure paracetamol for direct compression: the orthorhombic form. *Int J Pharm* 1996;**128**:1–8.
88. Nichols G, Frampton CS. Physicochemical characterization of the orthorhombic polymorph of paracetamol crystallized from solution. *J Pharm Sci* 1998;**87**:684–93.
89. Joiris E, Martino PD, Berneron C, Guyot-Hermann A-M, Guyot J-C. Compression behavior of orthorhombic paracetamol. *Pharm Res* 1998;**15**:1122–30.
90. Sun C, Grant DJW. Influence of crystal structure on the tableting properties of sulfamerazine polymorphs. *Pharm Res* 2001;**18**:274–80.
91. Sun C, Grant DJW. Improved tableting properties of p-hydroxybenzoic acid by water of crystallization: a molecular insight. *Pharm Res* 2004;**21**:382–6.
92. Zhang GGZ, Paspal SYL, Suryanarayanan R, Grant DJW. Racemic species of sodium ibuprofen: characterization and polymorphic relationships. *J Pharm Sci* 2003;**92**:1356–66.
93. Yada S, Ohya M, Ohuchi Y, Hamaura T, Wakiyama N, Usui F, et al. Solid phase transitionof CS-891 enantiotropes during grinding. *Int J Pharm* 2003;**255**:69–79.
94. Giordano F, Rossi A, Moyano JR, Gazzaniga A, Massarotti V, Bini M, et al. Polymorphism of rac-5,6-diisobutyryloxy-2-methylamino-1,2,3,4-tetrahydro-naphthalene hydrochloride (CHF 1035). I. Thermal, spectroscopic, and X-ray diffraction properties. *J Pharm Sci* 2001;**90**:1154–63.
95. Li Y, Han J, Zhang GGZ, Grant DJW, Suryanarayanan R. *In situ* dehydration of carbamazepine dihydrate: a novel technique to prepare amorphous anhydrous carbamazepine. *Pharm Dev Technol* 2000;**5**:257–66.
96. Han J, Suryanarayanan R. A method for the rapid evaluation of the physical stability of pharmaceutical hydrates. *Thermochim Acta* 1999;**329**:163–70.
97. Han J, Suryanarayanan R. Influence of environmental conditions on the kinetics and mechanism of dehydration of carbamazepine dehydrate. *Pharm Dev Technol* 1998;**3**:587–96.
98. Topley B, Smith ML. Function of water vapor in the dissociation of a salt hydrate. *Nature (London, United Kingdom)* 1931;**128**:302.
99. Smith ML, Topley B. Rate of dissociation of salt hydrates-reaction CuSO4.5H2O = CuSO4 H2O + 4H2O. *Proc R Soc London A* 1931;**134**:224–45.
100. Topley B, Smith ML. Kinetics of salt-hydrate dissociations: $MnC_2O_4,2H_2O = MnC_2O_4 + 2H_2O$. *J Chem Soc* 1935;321–5.
101. Garner WE, Jennings TJ. Nucleation phenomena arising during the dehydration of solid hydrates. *Proc R Soc London A* 1954;**224**:460–71.
102. Garner WE, Tanner MG. The dehydration of copper sulfate pentahydrate. *J Chem Soc* 1930;47–57.
103. Dollimore D, Jones TE, Spooner P. Thermal decomposition of oxalates. XI. Dehydration of calcium oxalate monohydrate. *J Chem Soc Sec A* 1970;**17**:2809–12.
104. Dollimore D, Heal GR, Mason J. The thermal decomposition of oxalates. Part 14. Dehydration of magnesium oxalate dihydrate. *Thermochim Acta* 1978;**24**:307–13.
105. Masuda Y, Nagagata K. The effect of water vapor pressure on the kinetics of the thermal dehydration of zinc formate dihydrate. *Thermochim Acta* 1989;**155**:255–61.
106. Masuda Y, Ito Y. The effect of water vapor pressure on the thermal dehydration of yttrium formate dihydrate. *J Therm Anal* 1992;**38**:1793–9.
107. Seto Y, Sato H, Masuda Y. Effect of water vapor pressure on thermal dehydration of lithium sulfate monohydrate. *Thermochim Acta* 2002;**388**:21–5.
108. Cardew PT, Davey RJ. The kinetics of solvent-mediated phase-transformations. *Proc R Soc London A* 1985;**398**:415–28.
109. Rodríguez-Hornedo N, Lechuga-Ballesteros D, Wu HJ. Phase transition and heterogeneous/epitaxial nucleation of hydrated and anhydrous theophylline crystals. *Int J Pharm* 1992;**85**:149–62.
110. Zhang GGZ, Gu C, Zell MT, Burkhardt RT, Munson EJ, Grant DJW. Crystallization and transitions of sulfamerazine polymorphs. *J Pharm Sci* 2002;**91**:1089–100.
111. Giordano F, Bettinetti GP, Caramella C, Conte U. Effects of grinding on the phase transitions of polymorphic modifications of sulfamethoxydiazine. *Bollettino Chimico Farmaceutico* 1977;**116**:433–8.
112. Cheng WT, Lin SY, Li MJ. Raman microspectroscopic mapping or thermal system used to investigate milling-induced solid-state conversion of famotidine polymorphs. *J Raman Spectrosc* 2007;**38**:1595–601.
113. Etter MC, Reutzel SM, Choo CG. Self-organization of adenine and thymine in the solid state. *J Am Chem Soc* 1993;**115**: 4411–12.
114. Kuroda R, Imai Y, Tajima N. Generation of a co-crystal phase with novel coloristic properties via solid state grinding procedures. *ChemComm* 2002;2848–9.
115. Duddu SP, Grant DJW. Formation of the racemic compound of ephedrine base from a physical mixture of its enantiomers in the solid, liquid, solution, or vapor state. *Pharm Res* 1992;**9**:1083–91.
116. Bucar DK, Henry RF, Lou X, Borchardt TB, Macgillivray LR, Zhang GGZ. Novel co-crystals of caffeine and hydroxy-2-naphthoic acids: unusual formation of the carboxylic acid dimer in the presence of a heterosynthon. *Mol Pharm* 2007;**4**:339–46.
117. Bucar DK, Henry RF, Lou X, Borchardt TB, Zhang GGZ. A "hidden" co-crystal of caffeine and adipic acid. *Chem Comm* 2007;525–7.
118. Serajuddin ATM. Solid dispersion of poorly water-soluble drugs: early promises, subsequent problems, and recent breakthroughs. *J Pharm Sci* 1999;**88**:1058–66.
119. Fox TG, Flory PJ. Second order transition temperatures and related properties of polystyrene. 1: influence of molecular weight. *J Appl Phys* 1950;**21**:581–91.
120. Fox TG, Flory PJ. Further studies on the melt viscosity of polyisobutylene. *J Phys Chem* 1951;**55**:221–34.
121. Fox TG, Flory PJ. The glass temperature and related properties of polystyrene—influence of molecular weight. *J Polym Sci* 1954;**14**:315–19.
122. Cohen MH, Turnbull D. Molecular transport in liquids and glasses. *J Chem Phys* 1959;**31**:1164–9.
123. Turnbull D, Cohen MH. Free-volume model of amorphous phase glass transitions. *J Chem Phys* 1961;**34**:120–5.
124. Turnbull D, Cohen MH. On free-volume model of liquid-glass transitions. *J Chem Phys* 1970;**52**:3038–41.
125. Doolittle AK. Studies in Newtonian flow. II. The dependence of the viscosity of liquids on free-space. *J Appl Phys* 1951; **22**:1471–5.
126. Gibbs JH, Dimarzio EA. Nature of the glass transition and the glassy state. *J Chem Phys* 1958;**28**:373–83.
127. Adam G, Gibbs JH. On the temperature dependence of cooperative relaxation properties in glass-forming liquids. *J Chem Phys* 1965;**43**:139–46.

128. Angell CA. Why C1 5 16–17 in the WLF equation is physical—and the fragility of polymers. *Polymer* 1997;**38**:6261–6.
129. Goldstein M. Viscous liquids and the glass transition IV. Thermodynamic equations and the transition. *J Phys Chem* 1973;**77**:667–73.
130. Johari GP. A resolution for the enigma of a liquid's configurational entropy-molecular kinetics relation. *J Chem Phys* 2000;**112**:8958–69.
131. Johari GP. The entropy loss on supercooling a liquid and anharmonic contributions. *J Chem Phys* 2002;**116**:2043–6.
132. Goldstein M. Viscous liquids and the glass transition. V. Sources of the excess specific heat of the liquid. *J Chem Phys* 1976;**64**:4767–74.
133. Hodge IM. Enthalpy relaxation and recovery in amorphous materials. *J Non-Cryst Solids* 1994;**169**:211–66.
134. Vogel H. Das Temperaturabhängigkeitsgesetz der Viskosität von Flüssigkeiten. *Phys Z* 1921;**22**:645–6.
135. Tammann VG, Hesse W. Die abhängigkeit der viscosität von der temperatur bei unterkühlten flüssigkeiten. *Z Anorg Allg Chem* 1926;**156**:245–57.
136. Fulcher GS. Analysis of recent measurements of the viscosity of glasses. *J Am Ceram Soc* 1925;**8**:339–55.
137. Angell CA, Smith DL. Test of the entropy basis of the Vogel-Tammann-Fulcher equation. Dielectric relaxation of polyalcohols near Tg. *J Phys Chem* 1982;**86**:3845–52.
138. Angell CA. Formation of glasses from liquids and biopolymers. *Science* 1995;**267**:1924–35.
139. Bohmer R, Ngai KL, Angell CA, Plazek DJ. Nonexponential relaxations in strong and fragile glass formers. *J Chem Phys* 1999;**99**:4201–9.
140. Moynihan CT, Easteal AJ, Wilder J. Dependence of the glass transition temperature on heating and cooling rate. *J Phys Chem* 1974;**78**:2673–7.
141. Moynihan CT, Lee SK, Tatsumisago M, Minami T. Estimation of activation energies for structural relaxation and viscous flow from DTA and DSC experiments. *Thermochim Acta* 1996;**280/281**:153–62.
142. Zhou D, Grant DJW, Zhang GGZ, Law D, Schmitt EA. A calorimetric investigation of thermodynamic and molecular mobility contributions to the physical stability of two pharmaceutical glasses. *J Pharm Sci* 2007;**96**:71–83.
143. Hodge IM. Effects of annealing and prior history on enthalpy relaxation in glassy polymers. 6. Adam-Gibbs formulation of nonlinearity. *Macromolecules* 1987;**20**:2897–908.
144. Ritland HN. Limitations of the fictive temperature concept. *J Am Ceram Soc* 1956;**39**:403–6.
145. Tool AQ. Relation between inelastic deformability and thermal expansion of glass in its annealing range. *J Am Ceram Soc* 1946;**29**:240–53.
146. Tool AQ, Eichlin CG. Variations caused in the heating curves of glass by heat treatment. *J Am Ceram Soc* 1931;**14**:276–308.
147. Narayanaswamy OS. A model of structural relaxation in glass. *J Am Ceram Soc* 1971;**54**:491–8.
148. Scherer GW. Use of the Adam-Gibbs equation in the analysis of structural relaxation. *J Am Ceram Soc* 1984;**67**:504–11.
149. Gordon M, Taylor JS. Ideal copolymers and the second-order transitions of synthetic rubbers 1: non-crystalline copolymers. *J Appl Chem* 1952;**2**:493–8.
150. Couchman PR, Karasz FE. A classical thermodynamic discussion of the effect of composition on glass-transition temperatures. *Macromolecules* 1978;**11**:117–19.
151. Nishi T, Wang TT. Melting-point depression and kinetic effects of cooling on crystallization in poly(vinylidene fluoride) poly(methyl methacrylate) mixtures. *Macromolecules* 1975;**8**:905–15.
152. Hoei Y, Yamaura K, Matsuzawa S. A lattice treatment of crystalline solvent-amorphous polymer mixtures on melting-point depression. *J Phys Chem* 1992;**96**:10584–6.
153. Marsac PJ, Shamblin SL, Taylor LS. Theoretical and practical approaches for prediction of drug-polymer miscibility and solubility. *Pharm Res* 2006;**23**:2417–26.
154. Zografi G. The amorphous state. In: *The 42nd Annual International Industrial Pharmaceutical Research and Development Conference*. Madison, WI; 2000.
155. Aso Y, Yoshioka S, Kojima S. Relationship between water mobility, measured as nuclear magnetic relaxation time, and the crystallization rate of amorphous nifedipine in the presence of some pharmaceutical excipients. *Chem Pharm Bull* 1996;**44**:1065–7.
156. Aso Y, Yoshioka S, Kojima S. Relationship between the crystallization rate of amorphous nifedipine, phenobarbital, and flopropione, and their molecular mobility as measured by their enthalpy relaxation and 1H NMR relaxation times. *J Pharm Sci* 2000;**89**:408–16.
157. Aso Y, Yoshioka S, Kojima S. Explanation of the crystallization rate of amorphous nifedipine and phenobarbital from their molecular mobility as measured by 13C nuclear magnetic resonance relaxation time and the relaxation time obtained from the heating rate dependence of the glass transition temperature. *J Pharm Sci* 2001;**90**:798–806.
158. Duddu SP, Zhang G, Dal Monte PR. The relationship between protein aggregation and molecular mobility below the glass transition temperature of lyophilized formulations containing a monoclonal antibody. *Pharm Res* 1997;**14**:596–600.
159. Guo Y, Byrn SR, Zografi G. Physical characteristics and chemical degradation of amorphous quinapril hydrochloride. *J Pharm Sci* 2000;**89**:128–43.
160. Yoshioka S, Aso Y, Kojima S. Temperature—and glass transition temperature—dependence of bimolecular reaction rates in lyophilized formulations described by the Adam-Gibbs-Vogel equation. *J Pharm Sci* 2004;**96**:1062–9.
161. Yoshioka S, Aso Y. A quantitative assessment of the significance of molecular mobility as a determinant for the stability of lyophilized insulin formulation. *Pharm Res* 2005;**22**:1358–64.
162. Yoshioka S, Aso Y, Miyazaki T. Negligible contribution of molecular mobility to the degradation rate of insulin lyophilized with poly(vinylpyrrolidone). *J Pharm Sci* 2006;**95**:939–43.
163. Hikima T, Adachi Y, Hanaya M, Oguni M. Determination of potentially homogeneous-nucleation-based crystallization in o-terphenyl and an interpretation of the nucleation-enhancement mechanism. *Phys Rev B* 1995;**52**:3900–8.
164. Wu T, Yu L. Origin of enhanced crystal growth kinetics near Tg probed with indomethacin polymorphs. *J Phys Chem B* 2006;**110**:15694–9.
165. Ishida H, Wu T, Yu L. Sudden rise of crystal growth rate of nifedipine near Tg without and with polyvinylpyrrolidone. *J Pharm Sci* 2007;**96**:1131–8.
166. Sun Y, Xi H, Ediger MD, Yu L. Diffusionless crystal growth from glass has precursor in equilibrium liquid. *J Phys Chem B* 2008;**112**:661–4.
167. Sun Y, Xi H, Chen S, Ediger MD, Yu L. Crystallization near glass transition: transition from diffusion-controlled to diffusionless crystal growth studied with seven polymorphs. *J Phys Chem* 2008;**112**:661–4.
168. Turnbull D, Fisher JC. Rate of nucleation in condensed systems. *J Chem Phys* 1949;**17**:71–3.
169. Hancock BC, Shamblin SL, Zografi G. Molecular mobility of amorphous pharmaceutical solids below their glass transition temperature. *Pharm Res* 1995;**12**:799–806.

170. Shamblin SL, Tang X, Chang L, Hancock BC, Pikal MJ. Characterization of the time scales of molecular motion in pharmaceutically important glasses. *J Phys Chem B* 1999;**103**:4113–21.
171. Zhou D, Zhang GGZ, Law D, Grant DJW, Schmitt EA. Physical stability of amorphous pharmaceuticals: importance of configurational thermodynamic quantities and molecular mobility. *J Pharm Sci* 2002;**91**:1863–72.
172. Zhou D. *Molecular mobility, physical stability, and transformation kinetics of amorphous and hydrated pharmaceuticals*. Ph.D. Minneapolis: University of Minnesota; 2003.
173. Ediger MD, Harrowell P, Yu L. Crystal growth kinetics exhibit a fragility-dependent decoupling from viscosity. *J Chem Phys* 2008;**128** 034709
174. Mehta SC. Mechanistic studies of linear single crystal growth rates of sulfathiazole and their inhibition by polyvinyl pyrrolidone. In: *Preparation and dissolution of high-energy sulfathiazole polyvinyl pyrrolidone coprecipitates*; 1969.
175. Simonelli AP, Mehta SC, Higuchi WI. Inhibition of sulfathiazole crystal growth by polyvinylpyrrolidone. *J Pharm Sci* 1970;**59**:633–8.
176. Sekikawa H, Nakano M, Arita T. Inhibitory effect of poly(vinylpyrrolidone) on the crystallization of drugs. *Chem Pharm Bull* 1978;**26**:118–26.
177. Ziller KH, Rupprecht H. Control of crystal growth in drug suspensions. 1) design of a control unit and application to acetaminophen suspensions. *Drug Dev Ind Pharm* 1988;**14**:2341–70.
178. Ma X, Taw J, Chiang C-M. Control of drug crystallization in transdermal matrix system. *Int J Pharm* 1996;**142**:115–19.
179. Ziller KH, Rupprecht H. Control of crystal growth in drug suspensions. III. Isothermal crystallization in the presence of polymers. *PZ Wissenschaft* 1990;**3**:147–52.
180. Wen H, Morris KR, Park K. Synergic effects of polymeric additives on dissolution and crystallization of acetaminophen. *Pharm Res* 2008;**25**:349–58.
181. Uekama K, Ikegami K, Wang Z, Horiuchi Y, Hirayama F. Inhibitory effect of 2-hydroxypropyl-b-cyclodextrin on crystal growth of nifedipine during storage: superior dissolution and oral bioavailability compared with poly(vinylpyrrolidone) K-30. *J Pharm Pharmacol* 1992;**44**:73–8.
182. Suzuki H, Sunada H. Influence of water-soluble polymers on the dissolution of nifedipine solid dispersions with combined carriers. *Chem Pharm Bull* 1998;**46**:482–7.
183. Matsumoto T, Zografi G. Physical properties of solid molecular dispersions of indomethacin with poly(vinylpyrrolidone) and poly(vinylpyrrolidone-co-vinyl-acetate) in relation to indomethacin crystallization. *Pharm Res* 1999;**16**:1722–8.
184. Crowley KJ, Zografi G. The effect of low concentrations of molecularly dispersed poly(vinylpyrrolidone) on indomethacin crystallization from the amorphous state. *Pharm Res* 2003;**20**:1417–22.
185. Khougaz K, Clas S-D. Crystallization inhibition in solid dispersions of MK-0591 and poly(vinylpyrrolidone) polymers. *J Pharm Sci* 2000;**89**:1325–34.
186. Huang J, Wigent RJ, Schwartz JB. Drug-polymer interaction and its significance on the physical stability of nifedipine amorphous dispersion in microparticles of an ammonio methacrylate copolymer and ethylcellulose binary blend. *J Pharm Sci* 2007;**97**:251–62.
187. Miyazaki T, Yoshioka S, Aso Y. Physical stability of amorphous acetanilide derivatives improved by polymer excipients. *Chem Pharm Bull* 2006;**54**:1207–10.
188. Konno H, Taylor LS. Influence of different polymers on the crystallization tendency of molecularly dispersed amorphous felodipine. *J Pharm Sci* 2006;**95**:2692–705.
189. Miyazaki T, Yoshioka S, Aso Y, Kojima S. Ability of polyvinylpyrrolidone and polyacrylic acid to inhibit the crystallization of amorphous acetaminophen. *J Pharm Sci* 2004;**93**:2710–17.
190. Kohri N, Yamayoshi Y, Xin H, Iseki K, Sato N, Todo S, et al. Improving the oral bioavailability of albendazole in rabbits by the solid dispersion technique. *J Pharm Pharmacol* 1999;**51**:159–64.
191. Zhang GGZ. *Influences of Solvents on properties, structures, and crystallization of pharmaceutical solids. Ph D Thesis*. Minneapolis, MN: University of Minnesota; 1998.
192. Miller JM, Collman BM, Greene LR, Grant DJW, Blackburn AC. Identifying the stable polymorph early in the drug discovery-development process. *Pharm Dev Technol* 2005;**10**:291–7.
193. Morissette SL, Almarsson O, Peterson ML, Remenar JF, Read MJ, Lemmo AV, et al. High-throughput crystallization: polymorphs, salts, co-crystals and solvates of pharmaceutical solids. *Adv Drug Del Rev* 2004;**56**:275–300.
194. Florence AJ, Johnston A, Price SL, Nowell H, Kennedy AR, Shankland N. An automated parallel crystallisation search for predicted crystal structures and packing motifs of carbamazepine. *J Pharm Sci* 2006;**95**:1918–30.
195. Remenar JF, Macphee JM, Larson BK, Tyagi VA, Ho JH, Mcilroy DA, et al. Salt selection and simultaneous polymorphism assessment via high-throughput crystallization: the case of sertraline. *Org Proc Res Dev* 2003;**7**:990–6.
196. Peterson ML, Morissette SL, Mcnulty C, Goldsweig A, Shaw P, Lequesne M, et al. Iterative high-throughput polymorphism studies on acetaminophen and an experimentally derived structure for Form III. *J Am Chem Soc* 2002;**124**:10958–9.
197. Chyall LJ, Tower JM, Coates DA, Houston TL, Childs SL. Polymorph generation in capillary spaces: the preparation and structural analysis of a metastable polymorph of nabumetone. *Cryst Grow Design* 2002;**2**:505–10.
198. Hilden JL, Reyes CE, Kelm MJ, Tan JS, Stowell JG, Morris KR. Capillary precipitation of a highly polymorphic organic compound. *Cryst Grow Design* 2003;**3**:921–6.
199. Childs SL, Chyall LJ, Dunlap JT, Coates DA, Stahly BC, Stahly GP. A Metastable polymorph of metformin hydrochloride: isolation and characterization using capillary crystallization and thermal microscopy techniques. *Cryst Grow Design* 2004;**4**:441–9.
200. Li L, Mustafi D, Fu Q, Tershko V, Chen DL, Tice JD, et al. Nanoliter microfluidic hybrid method for simultaneous screening and optimization validated with crystallization of membrane proteins. *Proc Natl Acad Sci* 2006;**103**:19243–8.
201. Gerdts CJ, Tereshko V, Yadav MK, Dementieva I, Collart F, Joachimiak A, et al. Time-controlled microfluidic seeding in nL-volume droplets to separate nucleation and growth stages of protein crystallization. *Angew Chem Int Ed Engl* 2006;**45**:8156–60.
202. Zheng B, Tice JD, Roach LS, Ismagilov RF. Crystal growth: a droplet-based, composite PDMS/glass capillary microfluidic system for evaluating protein crystallization conditions by microbatch and vapor-diffusion methods with on-chip X-ray diffraction. *Angew Chem Int Ed Engl* 2004;**43**:2508–11.
203. Talreja S, Kenis P JA, Zukoski CF. A kinetic model to simulate protein crystal growth in an evaporation-based crystallization platform. *Langmuir* 2007;**23**:4516–22.
204. He G, Bhamidi V, Tan RBH, Kenis PJA, Zukoski CF. Determination of critical supersaturation from microdroplet evaporation experiments. *Cryst Grow Design* 2006;**6**:1175–80.
205. He G, Bhamidi V, Wilson SR, Tan RBH, Kenis PJA, Zukoski CF. Direct growth of g -glycine from neutral aqueous solutions by slow, evaporation-driven crystallization. *Cryst Grow Design* 2006;**6**:1746–9.
206. Dombrowski RD, Litster JD, Wagner NJ, He Y. Crystallization of alpha-lactose monohydrate in a drop-based microfluidic crystallizer. *Chem Eng Sci* 2007;**62**:4802–10.

This page intentionally left blank

CHAPTER

3

Solid-State Characterization and Techniques

D. Law and D. Zhou

AbbVie Inc., North Chicago, IL, United States

3.1 INTRODUCTION

Understanding characterization of pharmaceutical solids encompasses three areas: (1) single-component characterization such as raw material, active pharmaceutical ingredient (API), excipient; (2) formulation or multicomponent characterization; and (3) understanding the pros and cons of the analytical techniques used in the characterization. All or most analytical techniques usually rely on measuring response to a perturbation. Therefore, analytical measurements of a solid pose a particular challenge because it is in a condensed state and thus has slower molecular mobility. Consequently for solids, be it crystalline or amorphous, the response is weaker than in a solution, and any additional heterogeneity contributes to variability in the signal. Heterogeneity arises from differences in particle size, morphology, trace impurities, and crystal defects that are inherent to all pharmaceutical solids. In addition, for formulations, the presence of excipients further complicates the analysis. Therefore, multiple methods are utilized before conclusions can be drawn. It is analogous to solving a jigsaw puzzle where each technique provides a piece of the answer, and at the end, each piece is required to sum up to provide a consistent picture. Consequently there exists a plethora of techniques that are utilized by pharmaceutical scientists, and these range from evaluating macroscopic properties such as surface area and particle size to molecular properties such as crystal packing, molecular mobility, and solid-state interaction. Additionally some of these techniques glean information about the powder bed, for example, flow or compressibility, while others provide information regarding intrinsic properties such as crystal packing. Furthermore, combination techniques, which are also called hyphenated techniques, for example, variable temperature-powder X-ray diffraction (PXRD), are becoming more and more popular in recent times because they allow multifaceted evaluation and increased sensitivity.

It is not possible to describe every technique a pharmaceutical scientist utilizes in the development of solid dosage form. Therefore, it is acknowledged that there are numerous commonly used techniques both macroscopic, such as specific surface area measurement, flow property evaluation, and compressibility, as well as molecular-level evaluations such as pycnometry, dielectric spectroscopy, and terahertz pulsed spectroscopy, that are not discussed in this chapter, nor is it possible to provide an in-depth description of each technique with respect to theory and fundamentals as well as applications. There are several excellent books available that describe the instrumental fundaments. The focus of the chapter will be to provide an understanding of some commonly used techniques with a few examples on the application of the techniques for pharmaceutical scientists.

The chapter will cover some of the more commonly used techniques that are used for solid-state characterization of both APIs as well as formulations. For each technique, the section provides an introduction to the technique including a brief summary of the fundamentals followed by examples/citations of the use of the technique for characterization for a single component (eg, API ingredient) as well as multicomponent (eg, formulation) system. The conclusion of each section will outline some pros and cons associated with each technique. The goal for each section will be to highlight the technique with respect to its application as well as to provide the reader with citations that are valuable for further research.

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00003-0

3.2 MICROSCOPY

Microscopy as a tool for analyzing pharmaceuticals is common because it forms a very powerful basis for characterization of pharmaceutical solids. Fundamentals of this technique can be found in several books.[1] There are several types of microscopes that are used for solid-state characterization, and a few have been discussed in this section. The method selected here considers the basis of interaction with the sample: (1) light or photon interaction is the basis for optical microscopy, (2) electrons instead of light constitute electron microscopy, and (3) a probe may be used to generate the image in the case of probe microscopy. The first two of the three techniques are most common; therefore, emphasis will be placed on optical and electron microscopy.

3.2.1 Optical microscopy

An optical microscope or light microscope uses visible light and a system of lenses to magnify a small sample. A basic optical microscope can be as simple as a magnifying glass (single lens microscope). For pharmaceutical applications, compound microscopes generally are used because they provide better resolution. Often these compound microscopes are connected to a charge-coupled device camera so that images can be directly viewed on a computer screen.

A cartoon of a compound optical microscope is presented in Fig. 3.1a. The ocular lens or the eyepiece provides 10× to 15× magnification and is connected via a barrel to the objective lenses. The objective lenses can be raised or lowered via the arm, and they revolve around the nosepiece. These lenses come in various magnifications, for example, 4×, 10×, 40×, and 100×, and are distinctly marked. Usually there are four lenses. Below the lens is the stage where the sample slide is placed. There are clips on this stage to hold the sample slide in place. The adjustment knobs on the side of the microscope involve both coarse and fine adjustments and are used to bring the sample into focus. Located below the stage is the lamp or the light source, and its brightness can be adjusted using the dimmer. The diaphragm is a circular disc between the lamp and the stage that varies the light intensity going through the sample.

A special case of compound microscope is the polarizing microscope, where a polarizer is placed after the light source. The sample is placed on a rotating stage, and an analyzer is placed before the eyepiece. Polarizing light microscopy exploits electromagnetic properties of light and anisotropy of crystalline solids to generate pleochroism.

If we follow Fig. 3.1b, light is an electromagnetic wave, or it is constructed of a longitudinal wave or the electric spectra, and at 90° to it is the lateral wave or the magnetic spectra. Focusing only on the electric wave vector, ordinary light or nonpolarized light has vectors at all different directions. When nonpolarized light is passed through an element called the polarizer, linear polarized light can be produced. This element has a slit that allows light to pass through in only in one direction. Crystals are usually anisotropic with intrinsic properties related to crystal packing, for example, refractive indices are different in different directions. There are, however, crystalline materials such as sodium chloride that are isotropic, but most pharmaceutical crystals are anisotropic. If an anisotropic material is placed in the path of the polarized light, it will split the beam into two perpendicular components: magnetic and electric. This ability to split the light is called birefringence. The velocities of the components will be dependent on the sample's intrinsic properties. When this light passes through the analyzer, which is simply another polarizer but placed perpendicular to the previous polarizer, the analyzer

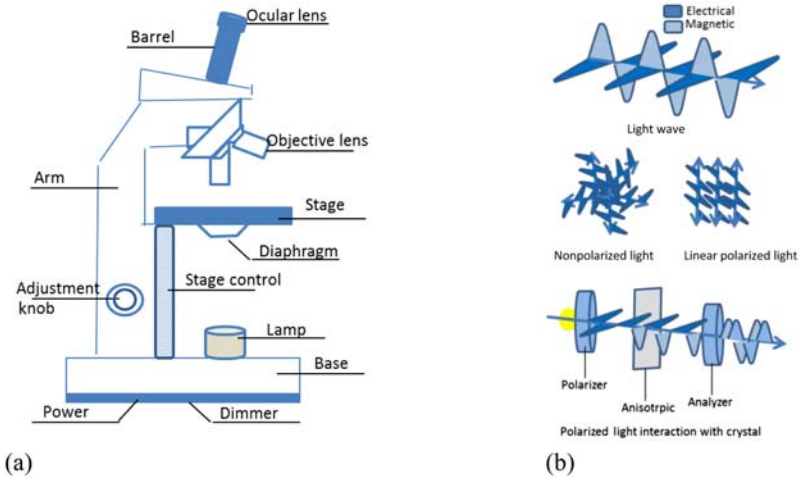


FIGURE 3.1 (a) Light microscope. (b) Polarized and unpolarized light and interaction of polarized light anisotropic sample.

combines the light. The difference in velocity will lead to constructive and destructive interference.

If the birefringent specimen is rotated, it produces brilliant colors as well as extinctions depending on the angle of the rotating stage. On the other hand, if the specimen is isotropic (ie, intrinsic properties are the same in all directions viewed with a set of polarizers), then either all light will pass through it and it looks bright, or no light will pass through, and it looks dark. Polarized light microscopy is a simple but a powerful tool to the pharmaceutical scientist and often is the first technique used to identify if the sample is amorphous or crystalline. In addition to optical microscopy and polarized light microscopy, there are other techniques such as fluorescent microscopy and phase contrast microscopy that are used to limited extent by pharmaceutical scientists.

Most pharmaceutical active ingredients are crystalline, and the molecules are packed to produce a 3-D periodicity over a long range. This 3-D periodicity influences a variety of properties, and one such influence is the way the material interacts with light. The spacing of molecules along the crystallographic axis determines the crystal system, most notably the optical properties. Extreme cases are diamond and graphite. In the case of diamond, the C-atoms atoms have a tetrahedral arrangement, creating planes that synergistically refract the light producing its brilliance, while the same C-atom assumes a planar structure with hexagonal packing that synergistically absorbs light, and it looks black. Likewise, pharmaceuticals do exhibit marketedly different optical properties and, therefore, can be easily identified using optical microscope. For example, ROY or 5-methyl-2-[(2-nitrophenyl)amino]-3-thiophenecarbonitrile, is known for its red, orange, and yellow crystals and has seven polymorphs with solved structures, which is the largest number in the Cambridge Structural Database (Fig. 3.2).[2]

Polarized light microscopes have frequently been used for monitoring, understanding, and observing phase changes. For example, the amorphous-to-crystalline transition[3] was monitored during the dissolution of indomethacin (Fig. 3.3a), while the degradation studies[4] of different crystal forms of indomethacin were monitored as seen in Fig. 3.3b.

As noted by the authors[2] themselves, it is rare in the pharmaceutical field to have polymorphs of different colors. Most active pharmaceuticals (both API and excipients) are white to off-white powders. However, when viewed under a polarized light microscope, they can exhibit birefringence that is inherent to the crystal form. Depending on pleochroism of the different polymorphs, it might be possible to use a polarized light microscope to identify different polymorphs. This method was particularly useful for studying phase transition of the low-dose pharmaceutical (Fig. 3.4).[5]

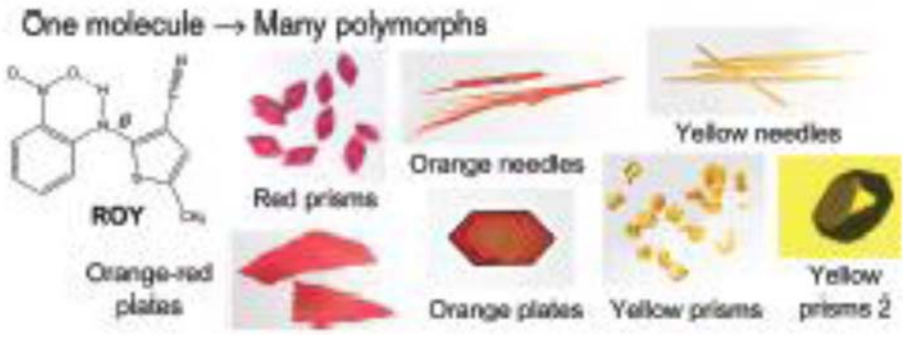


FIGURE 3.2 Crystalline polymorphs of ROY: different polymorphs viewed using optical microscope.

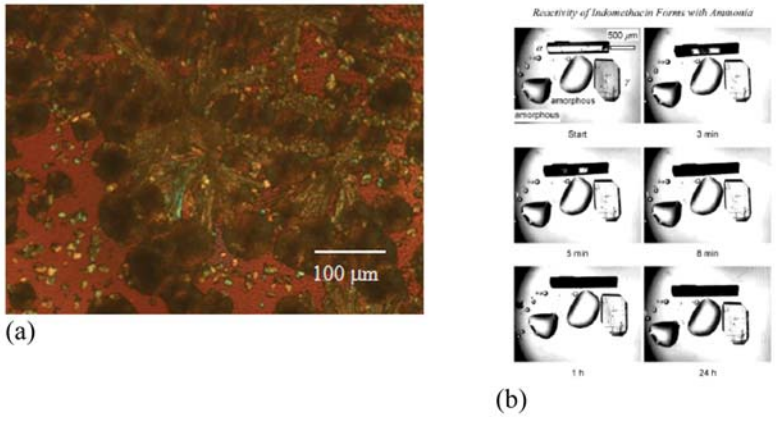


FIGURE 3.3 (a) Solvent-mediated crystallization from amorphous. (b) Decomposition of different forms of indomethacin.

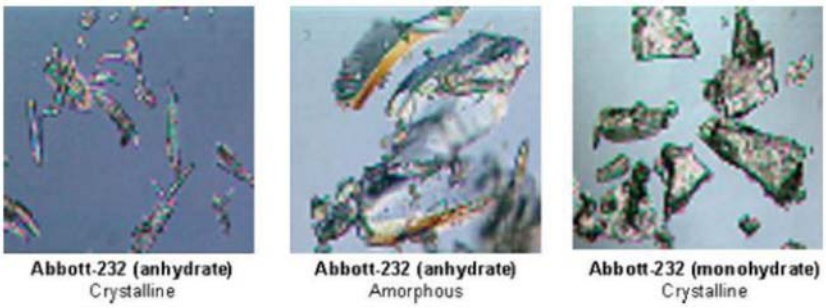


FIGURE 3.4 Different crystal forms exhibit different birefringence.

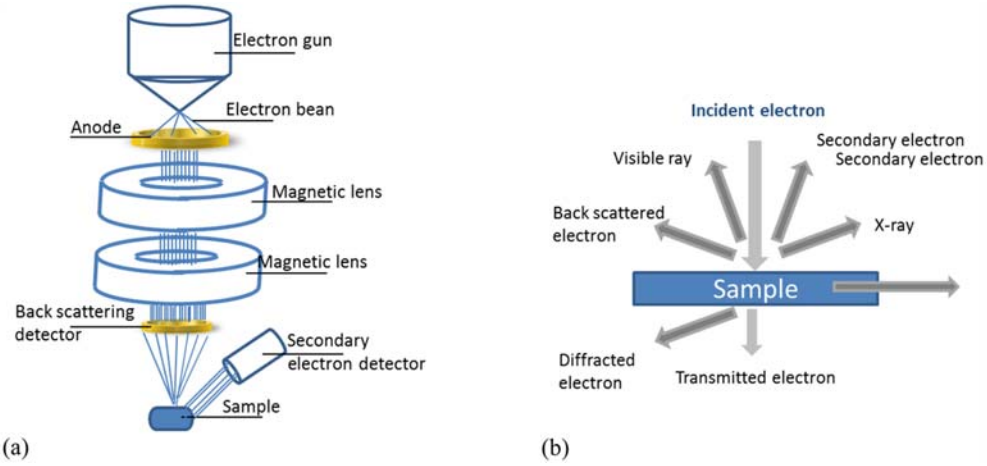


FIGURE 3.5 (a) Schematic principle of electron microscope. (b) Electrons scattering by the sample.

In this case, the authors found that when the crystals were oriented in a certain direction, one crystal form was pleochroic, while the other was not; thus they could identify the presence of the correct form in the tablet using a polarized light microscope. These are only a few applications of optical microscopy, illustrating that although it is a qualitative technique, it is also a useful tool for analyzing both API and formulations.

3.2.2 Electron microscopy

Unlike light microscopy, in the case of electron microscopy, electrons are used instead of light. There are two main types of electron microscopes: scanning electron microscopes (SEM) and transmission electron microscopes (TEM). As the name suggests, in SEM, the electron beam scans the surface, while in TEM the beam is transmitted through a thin specimen. For pharmaceutical solids, SEM is more widely used than TEM because the latter does require the production of thin films. Therefore, this section will focus on fundamentals and the use of SEM for pharmaceutical applications.

Fig. 3.5a provides a schematic for the operation of a SEM. It consists of an electron gun that releases electrons. This gun is usually a tungsten filament heated to 2400°C when it starts emitting electrons. These electrons go through an anode, which being positively charged accelerates and collimates the electrons. The accelerating potential can be varied by the operator from 1 to 30 kV. In the case of an optical microscope, glass lenses are used, while an electron microscope uses magnetic lens. The magnetic lens focuses the electrons into a narrow electron beam as well as controls the amount of electrons hitting the surface of the sample. Then there are detectors that collect the signal and are connected to computers that construct the image. Before discussing the detectors, it is important to take a minute to discuss what happens when the electron beam encounters the sample.

When electron beams encounter the sample, electrons can be absorbed, reflected, excite the surface releasing secondary electrons, or emit specific radiation (Fig. 3.5b). The reflected electron is the backscattered electrons, so there is always a backscattering detector. The most important signal comes from the secondary electrons. This signal is used to construct the surface structure of the sample. Since there is no light that is used for this technique, one cannot see the sample or colors; it is the distance these secondary electrons travel that allows for the construction of the image. There may be some sample-specific X-ray radiation generated by the absorbed electrons, so

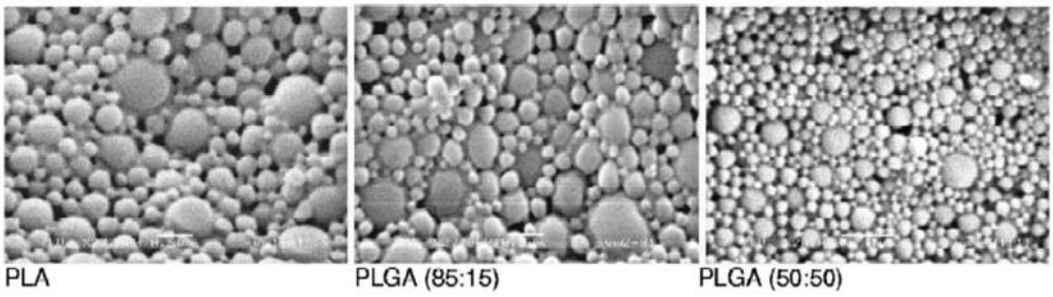


FIGURE 3.6 SEM image of TPGS impact on PLGA nanoparticles.

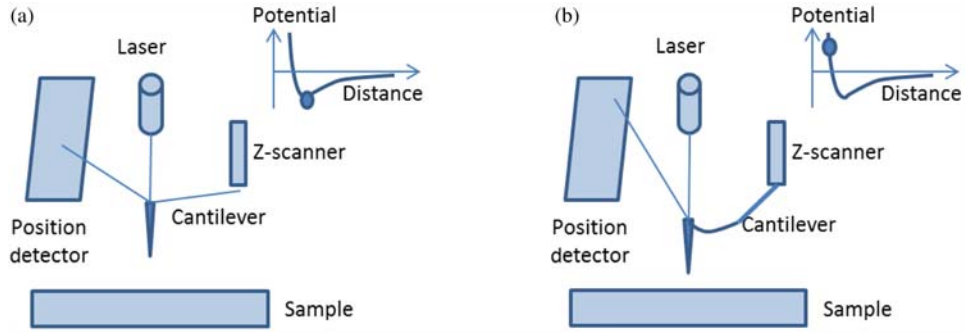


FIGURE 3.7 Schematic principle of an atomic force microscope with cantilever at a reference position (a) and with cantilever scanning down to sample surface (b).

sometimes SEMs have X-ray detectors in addition to backscattering and secondary electron detectors. These X-ray detectors, if present, can provide some information regarding the surface composition.

The most popular application of SEM in the field of pharmaceuticals is in the evaluation and understanding of the morphology of APIs, excipients, and granulations. Here is one application[6]: paclitaxel is an anticancer agent that has exhibited solubility-related drug delivery challenges. Although formulations with cremophor have proven to be successful in paclitaxel delivery, this surfactant has toxic side effects; therefore, TPGS-PLGA nanoparticles were explored to minimize toxicity. As seen in Fig. 3.6, SEM provided valuable information regarding the impact of nanoparticle morphology as a function of TPGS concentration as well as the mechanism in which the TPGS was incorporated.

3.2.3 Probe microscopy

Probe microscopy is the next step or an extension of electron microscopy. These methods allow still higher resolution than SEM or TEM. In this case, the sample surface at an atomic or molecular scale or nanoscale dimensions can be evaluated using a sharp cantilever tip. There are two types of microscopes in this field: (1) a scanning tunneling microscope (STM) and (2) an atomic force microscope (AFM). Since the AFM is more widely used in the pharmaceutical field, this section will focus only on AFM techniques and applications. The basic principle of AFM is outlined in Fig. 3.7. In AFM, a cantilever with a very sharp tip is used to scan the sample surface. As the cantilever approaches the surface, the attractive forces cause the cantilever to bend, but as the cantilever gets close to the surface such that it makes contact, repulsive forces takeover, causing the cantilever to deflect away from the surface. A laser coupled with a photo-diode is used to monitor the deflection of the cantilever. This deflection as a function of surface topography is then used to generate the sample image. This is the description when the instrument is operated in contact mode. AFM is currently available in contact mode, tapping mode, and peak force tapping mode, and the different modes are used to obtain the desired resolution.

Some of the early pioneering work using molecular crystals in the field of AFM evaluated crystallization from a solution phase. Fig. 3.8a shows the crystallization of calcite in the absence and presence of aspartic acid.[7] The addition of aspartic acid to growth solutions has a dramatic effect on the spiral morphology, with the crystal growth hillocks becoming asymmetric and rounded. The authors evaluated growth mechanisms in both organic solution as well as biomaterials, which led to greater understanding of the molecular

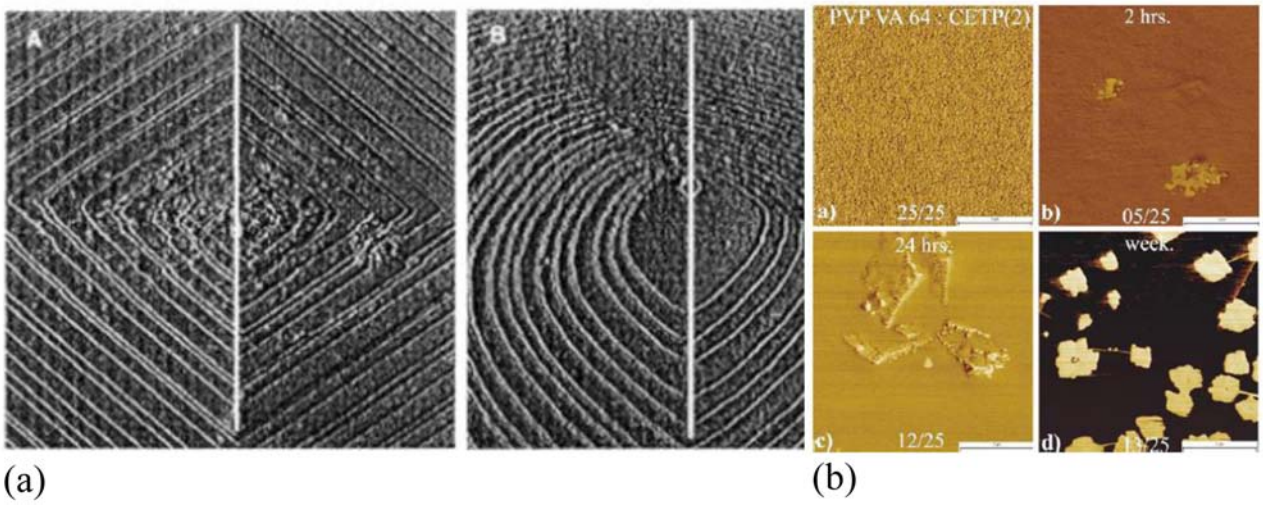


FIGURE 3.8 (a) Crystallization of pure calcite in the presence of aspartic acid. (b) Crystallization of amorphous formulation when exposed to high relative humidity.

organization in the crystal lattice. More recently, AFM was used to evaluate formulation stability. Again, an example has been provided[8]: In Fig. 3.8b, amorphous formulations (of CETP inhibitor) in a polymer (PVP VA 64) were exposed to 75% RH, and crystal growth was monitored as a function of time. The authors evaluated a list of polymers for developing amorphous formulation (Eudragit L100, HPMCAS, PVP, etc.), and these studies allowed them to choose the polymer that provided the best stability when exposed to humidity. These microscopic techniques, although not quantitative, are powerful in detecting small changes in samples with the best precision.

The microscopes described in this section vary from those that are simple and provide information regarding morphology to highly sophisticated ones that allow visualization of crystal growth steps. The key advantages of microscopy are that one can see the sample, small changes can be detected, and a very small amount of sample is required, but the technique is qualitative and visual, not quantitative.

3.3 POWDER X-RAY DIFFRACTION

X-ray diffraction techniques have been used for determining crystal structure, but in the pharmaceutical industry, by far the most common use of X-ray diffraction is associated with analyzing powder samples. This technique is nondestructive and is typically used for evaluating the degree of crystallinity, characterizing polymorphs and solvates, and the study of phase transitions.

Crystalline solids have long-range 3-D periodicity. The unit cell as shown in Fig. 3.9a is the basic repeating unit or symmetry that defines the crystal lattice structure. This unit repeats in all three dimensions over a long range to produce the crystal structure (Fig. 3.9b). It can be envisioned that the molecules in the unit cell can pack in different ways, producing different crystal structures. When a pharmaceutical solid can exist in different molecular packing, it is said to exhibit polymorphism. Since this crystal structure influences the intrinsic properties of a solid, understanding it is an integral aspect of drug development. A dramatic impact of packing can be demonstrated with diamond (Fig. 3.9c) and graphite (Fig. 3.9d). In diamond the tetrahedral packing of carbon results in planes that reflect light synergistically, producing its luster. In graphite the same carbon packs in laminar layers, allowing synergistic light absorption causing a black appearance. Diamond is one of the hardest known crystals, whereas in the case of graphite, the planar layers can slide under stress, making graphite a relatively soft crystal that can readily deform. This example demonstrates the consequence of packing the results in dramatically different responses to light or mechanical stress perturbation. In fact, crystal packing can influence thermodynamic properties such as melting and solubility, kinetic properties such as stability and dissolution, and surface properties such as morphology and spectroscopic properties.[9]

Therefore, in order to have consistent bioavailability, stability, or manufacturability, it is essential to maintain the desired crystal form for the API or excipient, thus making utilization of powder diffraction technique essential for drug development as well as on-market support. A brief description of the concept and experimental technique is discussed, but the reader should refer to specialized textbooks for details.[10,11]

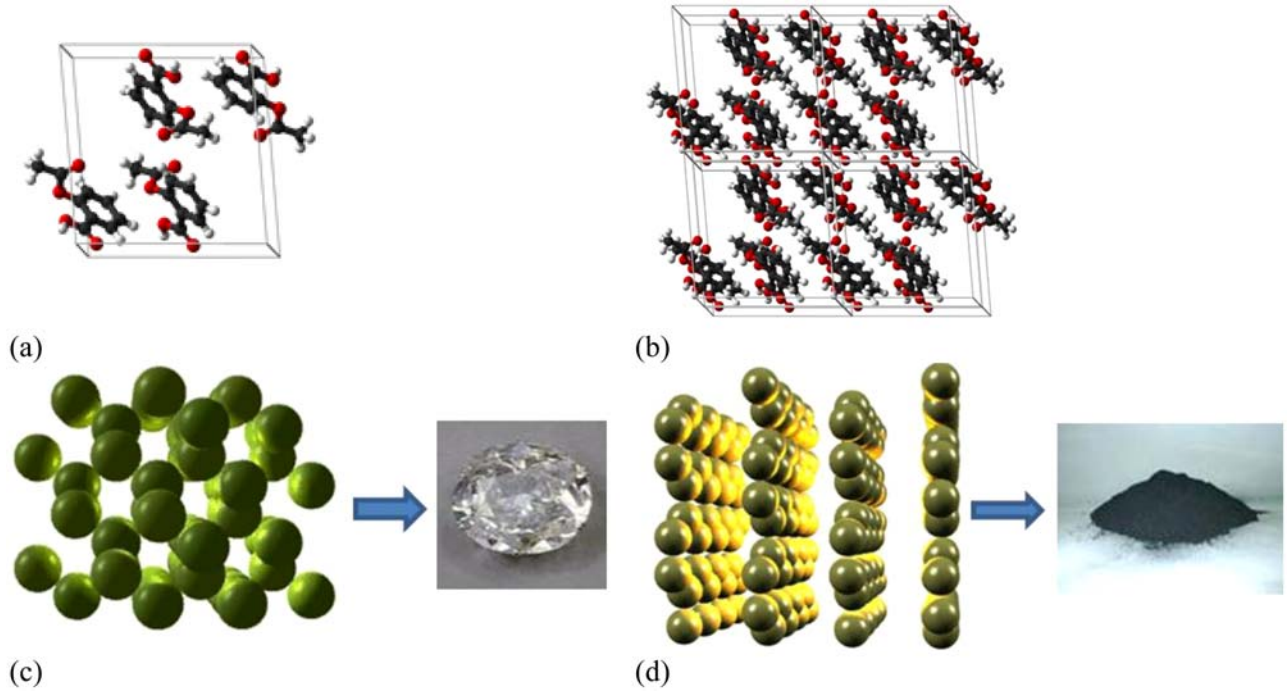


FIGURE 3.9 (a) Unit cell. (b) 3-D repeat of unit cell leads to the long-range periodicity. (c) Tetrahedral packing of C in diamond. (d) Planar packing of C in graphite.

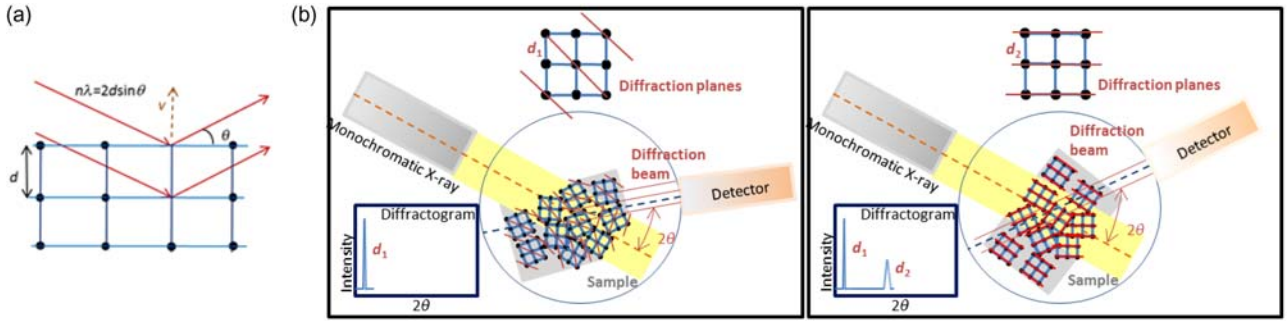


FIGURE 3.10 (a) Bragg diffraction condition for coherent scattering. (b) A θ:2θ arrangement for generating diffraction patterns.

The diffraction patterns depend on a series of planes in a crystal lattice. The orientation of these planes is a function of the molecular packing and is defined by Miller indices by considering how the planes intersect the main crystallographic axes of the solid. A set of rules leads to the assignment of the Miller indices (*hkl*), which is a group of numbers that represent the reciprocal axial intercepts and uniquely identify the planes of a crystal. Although Miller notations are used to define direction as well as planes, for the purpose of this discussion, only planes (*hkl*) and family of planes {*hkl*} will be considered.

Diffraction is a *scattering* phenomenon. X-rays are electromagnetic radiation that lie between ultraviolet (UV) and gamma radiation in the electromagnetic spectrum, and when X-rays are incidental to crystalline solids, they are scattered in all directions. Bragg's law (Fig. 3.10a) describes the conditions under which these scattered beams are completely in phase and reinforce one another to form diffracted beams. Thus, per Bragg's law, diffraction will occur if a perfectly parallel and monochromatic X-ray beam of wavelength λ is incident on a crystalline sample at an angle θ:

$$n\lambda = 2d_{hkl}\sin\theta$$

where d_{hkl} is the vector extending from the origin to the plane (*hkl*). The magnitude of this vector is the distance between parallel planes in the {*hkl*}. In most diffractometers, X-ray wavelength is fixed so the family planes produce peaks as a function θ (Fig. 3.10b).

The essential parts of an X-ray diffractometer are: (1) X-ray tube or the source; (2) incident-beam optics, which consists of the slits and monochromators that remove unwanted radiation, reduce divergence, and refocus the beam to parallel path; (3) the goniometer consisting of the platform that holds and moves the sample, optics, detector, and/or tube; (4) the sample and the sample holder; (5) the receiving-side optics that conditions the diffracted beam; and (6) the detector that counts the number of X-rays scattered by the sample.

Typically the source where X-rays are produced is an evacuated tube similar to cathode ray tubes. An applied current heats up a cathode filament (eg, tungsten), which liberates electrons. The liberated electrons are accelerated by a high-voltage potential toward a target, which is a metal anode. The bombarding electrons generate X-rays from the anode. Most instruments today use copper (Cu) as the anode. The bombarding electrons eject electrons from the inner shells of the atoms of the target metal (Cu) producing vacancies. Those vacancies are quickly filled by electrons dropping down from higher levels, emitting X-rays with sharply defined frequencies. The X-rays pass through a slit, which determines the angular width of the beams. Wider slits give more energy but also give wider diffraction lines and, therefore, lesser resolution. The beam then passes through a filter that removes undesired radiation. Note in the case of the Cu anode there are three types of radiation: CuKα1, CuKα2, and CuKβ. The lines for CuKα1 and CuKα2 are overlapping; these are used for most X-ray diffraction techniques. The low-energy CuKβ is filtered out by a nickel filter. Some very expensive instruments can also remove the CuKα2, providing extremely high resolution. The beams are incident to a sample at an angle θ, and the diffraction vector v bisects the incident and diffracted beam. X-rays at an angle θ are reflected by internal crystal planes separated by distance d. Bragg diffraction results from constructive interference when the quantity $2d\sin\theta$ is an integral number of wavelengths. Before discussing the sample or the detector, it is important discuss the goniometer.

The goniometers satisfy the Bragg–Brentano geometries where the diffraction vector v is normal to the sample. There are two different setups that are common. The setup shown in Fig. 3.10b is a θ:2θ arrangement where the source is fixed. The sample moves at θ°/minutes, while the detector rotates at $2\theta^\circ$/minutes. This setup allows for precision movement of the goniometer but has the disadvantage that the sample is vertical and often falls off the holder. In θ:θ arrangement, the sample is always vertical, while both the source and the detector move at θ°/minutes.

In Fig. 3.10b, the polycrystalline powder packed in a sample holder contains thousands of crystallites, allowing for the observance of all possible diffraction peaks. The basic assumption is that for every set of planes there is an equal number of crystallites that will diffract, and there is a statistically relevant number of crystallites. In reality only a small fraction of crystallites in the sample contribute to the diffraction pattern; therefore, unusably, the sensitivity is increased by spinning the sample. The diffraction patterns are collected as 2θ versus intensity, and for pharmaceuticals, this is how the data are represented. The more accurate representation would be to transform the data such that the representation eliminates instrumental parameters. This transformation involves normalizing the peak intensities to the area of the largest peak and using Bragg's relationship to convert the 2θ values to d_{hkl}, an intrinsic property since it corresponds to a family of planes in the sample.

Finally a few words regarding the detector: the two common types of detectors are a point detector and a position sensitive detector. One type of detector from each of these two categories will be briefly mentioned. The scintillation counter is a point detector that converts X-ray photons to an electrical signal in a two-stage process. The photons collide with a scintillator producing photons in the blue region of the visible spectrum. These photons are then converted to a signal or voltage pulse by means of a photomultiplier. The scintillation counters have high efficiency in response time, making them ideal for point intensity measurements, but the energy resolution is poor. Although gas-filled point detectors are available, these are less common than gas-filled position sensitive detectors. The gas-filled detectors work on the principle that X-ray photons can ionize inert gas. Since the ionization energy is low, an X-ray photon can produce several hundred ion pairs. These ion pairs are accelerated toward a wire causing further ionization. The burst of electrons is then converted to a voltage pulse, which is shaped and counted by the electronics. The use of special types of wires and a multichannel analyzer enables this type of detector to record data over a large range of scattering angles, improving the speed of acquisition.

PXRD is the gold standard for solid form identification. Today every drug development program starts by identifying the most developable crystal form and uses X-ray diffraction as a tool for characterizing API or the drug product. Therefore, there are various review articles written on this technique[12–14] as well as books.[15] Not only is the technique used for the identification of the solid phase, it has been used for studying the kinetics of phase transformation examples during processing,[16,17] quantitative determination of polymorphic phases in a mixture or quantitative determination of degree of crystallinity,[18] stability studies for amorphous formulations,[19] and the analysis of counterfeit product.[20]

This gold standard technique for phase identification is largely a qualitative technique. Through careful experimental design, quantitative information has been extracted, but it is not possible to obtain thermodynamic quantities or even information regarding functional groups in the molecules using this method.

3.4 THERMAL ANALYSIS

Thermal properties (eg, heat, heat capacity, mechanical, and gravimetric) accompanying physical and/or chemical changes may be revealed when a sample is subjected to a temperature program (heating, cooling, or isothermal). The methods for characterizing these changes are called thermal analysis, and they provide information on the underlying physical or chemical processes such as phase transition, chemical reaction, and decomposition. Since 1887 when Le Chatelier conducted the first thermal analysis experiment, the techniques have matured in theory, instrumentation, and automation, and they have become the standard tools used in many fields. Thermal analysis is widely used in characterizing pharmaceutical solids. A number of reviews are available on the applications of thermal analysis to pharmaceuticals.[21–24]

3.4.1 Differential scanning calorimetry

3.4.1.1 Instrumentation

Differential scanning calorimetry (DSC) is a further development of the earlier technique called differential thermal analysis (DTA). In both DSC and DTA, a sample is measured against an inert reference. In DTA, the sample and the reference are heated simultaneously in a single furnace with a symmetrical design. The temperatures of both parts are expected to be the same after a short equilibration time, but that of the sample may change at a different rate from the reference during a thermal event, such as freezing, melting, evaporation, dehydration/desolvation, or glass transition. The temperature difference (ΔT) between the sample and the reference is recorded during a controlled temperature program (heating/cooling). However, the classical DTA instrumentation design where thermocouple junction is inserted into the sample prevents a true quantitative analysis. Significant temperature inhomogeneity or temperature gradient may exist in the sample for classical DTA, which renders the quantitative results of a thermal event dependent on its apparent thermal resistance, which is a function of the nature of the material, morphology, and sample preparation such as packing density.

The Boersma design[25] introduced a controlled heat leak by inserting a low-resistance heat flow path between the sample and the reference by placing the thermocouple junction outside (just below the sample platform) the sample to reduce the impact of thermal resistance on ΔT. The Boersma design results in a much smaller ΔT because of the good thermal contact, which diminishes the impact on the measured heat flow by the apparent sample thermal resistance and makes the technique quantitative for calorimetric information. So the Boersma DTA becomes a calorimetry and bears the name of heat-flux DSC, or hf-DSC. The hf-DSC measures heat flow indirectly as a result of the ΔT, and the calorimetric information is obtained through calibration using a known standard.

Another design is called power compensation DSC (pc-DSC). In pc-DSC the sample and references are enclosed into separate furnaces that are controlled independently. The temperatures of the sample and the reference are made identical by varying the power input into the two furnaces. The difference in power measures the changes in enthalpy or heat capacity of the sample relative to the reference. Pc-DSC directly measures the enthalpy changes against a reference and has improved sensitivity.

Fig. 3.11 illustrates the key design differences between heat flux and power compensation DSCs.

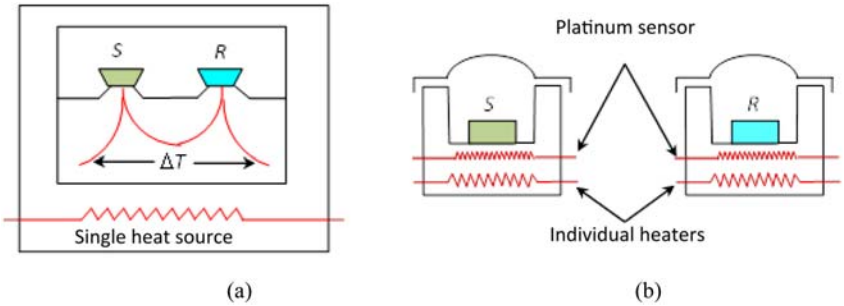


FIGURE 3.11 Schematic representation of hf-DSC (a) and pc-DSC (b) (*S* = sample, *R* = reference).

Conventional DSC employs a linear heating or linear cooling method, so the reference temperature profile may be represented as:

$$T = T_0 + ßt$$

where T_0 is the initial temperature, $ß$ is the linear heating rate (°C/minutes or K/minutes), and t is the heating time (minutes). The heat flow, dQ/dt, at constant pressure may be represented as:

$$\left(\frac{dQ}{dt}\right)_p = mC_p ß + f(T,t)$$

where m is the mass of the sample, C_p is the specific thermodynamic heat capacity ($J.K^{-1}.mol^{-1}$), and $f(T,t)$ represents heat flow due to a kinetic process (physical or chemical). The total heat flow is composed of two parts. One part is a function of the sample's true heat capacity, and the other, $f(T,t)$, is a function of temperature and time. However, these two components are not separated by the conventional DSC.

In temperature-modulated DSC (TMDSC), a periodic temperature modulation is added to a linear heating/cooling or isothermal temperature program. Temperature modulations such as sinusoids (sine or cosine) are the most convenient forms, but other forms such as sawtooth have also been used. The temperature program for a sinusoidal modulation may be represented as:

$$T = T_0 + ßt + A\sin(\omega t + \varphi)$$

where A is the amplitude, ω is the modulation frequency, and φ is phase angle.

Fourier transformation is applied in TMDSC to deconvolute the apparent heat flow signal into a reversible and a nonreversible component. The separation of heat flow facilitates the study and interpretation of many complex thermal events encountered for various materials. It also allows the thermodynamic heat capacity of a material to be measured more accurately. Fig. 3.12 provides an example of temperature modulation.

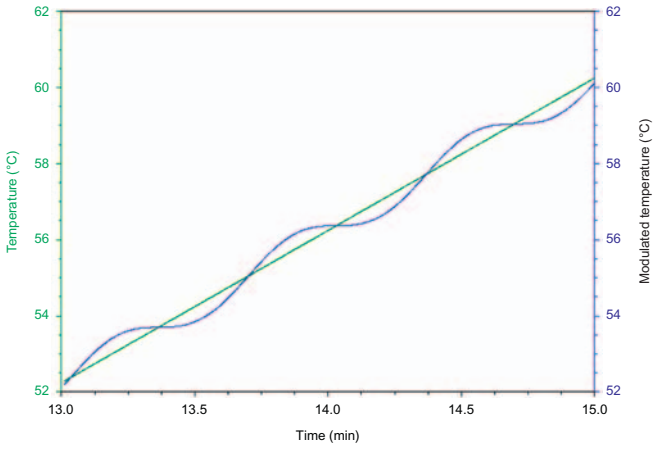


FIGURE 3.12 A representative temperature profile for TMDSC.

3.4.1.2 Applications

The following four applications are worth mentioning:

1. Melting and phase diagram.
2. Characterization of polymorphs.
3. Characterization of solvates/hydrates.
4. Characterization of amorphous phases.

3.4.1.2.1 Melting and phase diagram

DSC is widely used in the characterization of pharmaceutical solids commonly used in formulation/dosage form development. The melting temperature (T_m), enthalpy of melting (ΔH_m), and entropy of melting (ΔS_m) of a crystalline or partially crystalline material can be obtained readily from a simple DSC scan. As for any other first-order phase transitions, the change in free energy is zero during melting, therefore, $\Delta S_m = \Delta H_m/T_m$, where T_m is expressed on absolute scale (K). This relationship does not hold when melting is accompanied by other thermal events, such as decomposition, when there is no clean separation between those events.

Melting temperature depression occurs in the presence of impurities that are dissolvable and miscible in the liquid state. The formation of a liquid mixture lowers the free energy of the system; therefore, the system will melt at a lower temperature. The relationship can be derived based on ideal solution as follows:

$$\left(\frac{\partial \ln x_2}{\partial(1/T)}\right)_p = -\frac{\Delta H_m}{R}$$

$$\ln x = \frac{\Delta H_m}{R}\left(\frac{1}{T_{m,0}} - \frac{1}{T_m}\right)$$

where x is the concentration (molar fraction) of the substance of interest in the testing specimen, R is the universal gas constant, and $T_{m,0}$ is the melting temperature of the pure phase or a purity reference. This equation forms the basis of purity test by DSC.

DSC is a convenient tool to construct such phase diagrams of solids that are miscible in molten state. Mixtures of a binary system with varying compositions are scanned by DSC. The onset melting temperature of each component is recorded and then is plotted against the composition.

In a binary system consisting of two crystalline materials, the melting temperature of both will be lowered provided that the two materials are miscible in the liquid state, as dictated by the free energy decrease.

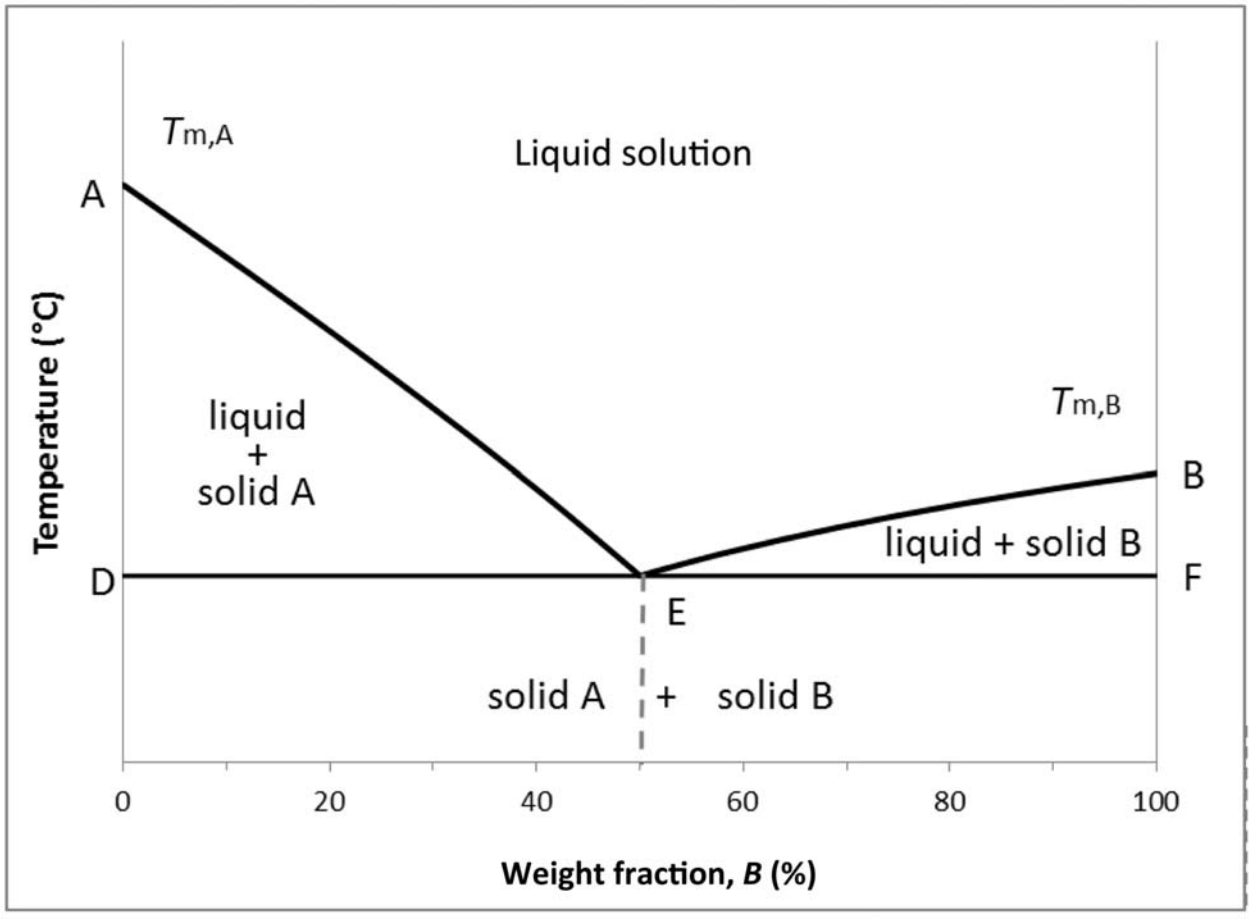


FIGURE 3.13 A basic phase diagram with eutectic melting.

Fig. 3.13 is a basic temperature-composition phase diagram of such a system. The melting temperature of component A follows the curve AE, which may be described by the above equation for melting depression of component A. Similarly, the melting point of component B follows the curve BE. At point E, both components reach equilibrium and melt at the same temperature. This temperature is called the eutectic temperature, and the corresponding composition is said to be a eutectic mixture.

3.4.1.2.2 Characterization of polymorphs

In the previous section on PXRD, it was discussed that polymorphism is the existence of more than one crystal packaging by the same molecule entity. Organic molecules including drugs are notorious for their formation of multiple crystalline forms.[26,27] Polymorphs can have a significant difference in solubility, stability, mechanical properties, and thermal properties and thus may impact formulation, manufacture, and performance of pharmaceutical products. Hence, agencies generally require that crystal forms of a drug substance be examined carefully against the properties that may be relevant to the dosage form design, process development, manufacturing, and in vivo performance of the product and implement appropriate controls to safeguard the identity, strength, quality, purity, and potency of the drug product.

Polymorphs represent the local minimums in the free energy landscape of the crystalline molecule. An important topic in polymorph characterization is to delineate the thermodynamic relationship among the different crystal forms. A thermodynamically more stable form is often preferred due to the less likelihood of phase transformation in general. Metastable forms could be used, and in some cases, they may be beneficial to the manufacturing process or in vivo performance. However, the incorporation of a metastable form may have significant implications to the dosage form design, manufacturing, storage, and regulatory compliance.

Two polymorphs may be thermodynamically related as enantiotropes or monotropes. In the temperature domain, an enantiotropic relationship exists if one form is more stable under one temperature range and the other more stable under another, that is, their conversion is thermodynamically reversible. On the

contrary, if one polymorph is always more stable at all temperatures (below the melting temperature), they are said to be related monotropically. The relationship could also be explained in terms of the transition temperature, T_t, which is the temperature where the two forms are in thermodynamic equilibrium (or change in free energy, $\Delta G = 0$). Two polymorphs are related enantiotropically if T_t lies below the melting temperatures of both forms; otherwise, they are monotropically related. The free energy landscape between any two polymorphs is represented as in Fig. 2.2.

The free energy relationship between two polymorphs may be determined by DSC. Polymorphic transformation is generally a solid-state phase transition and is often accompanied by a small change in enthalpy. If the interconversion is observed between two forms reversibly (ie, during heating and reversely during cooling), then the relationship is the enantiotropic. However, a transformation may not be observed in some cases due to insufficient kinetic rate, even if the transformation is thermodynamically favorable. In other words, a thermodynamically favorable polymorphic transformation may not happen during a DSC experiment. If a polymorphic transformation does occur at temperature T_{obs} during heating, the true transition temperature is somewhat below T_{obs} due to the kinetic lag. The same is true during a cooling cycle.

Burger and Ramburger[28,29] formulated several rules of polymorphic relationships based on DSC data. The heat of transition rule states that two forms are related enantiotropically if an endothermic transition is observed at some temperature, and they are related monotropically if an exothermic transition is observed at some temperature. The heat of fusion rule states that if the higher melting form has the lower heat of fusion, the two forms are usually enantiotropic. Otherwise they are monotropic. The form that melts at a higher temperature is always more stable at a high temperature. These two rules are well illustrated in Fig. 2.4.

Similarly to other rules of thumb, exceptions to the Burger and Ramburger rules do occur. A more comprehensive thermodynamic treatment of the polymorphic relationships was later introduced.[30] Yu's treatment considers both the fusion data and the temperature dependence of the extrapolated free energy differences and, therefore, can resolve the misclassifications by the Burger and Ramburger rules in some exceptions.

3.4.1.2.3 Characterization of hydrates

Many drug molecules form hydrates. A hydrate is a crystal packing fulfilled by the drug (host) molecules and water (guest) molecules via various intermolecular interactions, with hydrogen bonding (H-bond) interaction being particularly important. Nonspecific van der Waals interaction may also play important role in nonstoichiometric hydrates where the water molecules are simply occupying the void space in a crystal. Incorporation of water molecules in a crystal lattice changes many pharmaceutical properties.[31] A hydrate usually has lower aqueous solubility than its anhydrate, and a similar relationship exists between higher hydrates and lower hydrates. Other properties such as density, mechanical properties, physical and chemical stability could also be impacted.

Similarly to polymorphs (to that end a hydrate is also called a pseudopolymorph), physical stability of pharmaceutical hydrates is important to drug product development. Whether an anhydrate or a hydrate is selected for development, the risks of hydration/dehydration during manufacturing and storage should be understood because humidity, or the lack of, is ubiquitous in our environment. Appropriate measures should be implemented to prevent the inadvertent form conversion.

DSC is routinely used to characterize hydrates. As indicated in Chapter 2 "Crystalline and Amorphous Solids," a critical relative humidity, RH_C, exists for each hydrate. A hydrate is more stable at $RH > RH_C$ than the anhydrate or the lower hydrate, if it exists, and vice versa. While a conventional DSC is not capable of determining the critical RH for a hydrate, it is used extensively to identify and characterize a hydrate, and it provides useful information when combined with other complementary tools such as gravimetric analysis, Karl Fischer titrimetry, mass spectroscopy, and Fourier-transformed infrared (FTIR). Dehydration is often revealed as endothermic peaks during a DSC scan. Dehydration often occurs at lower temperatures (eg, $<100^\circ C$), reflecting the relatively weak host-guest interactions. A relatively weak hydrate may start to lose water immediately after being introduced in the DSC cell when purged with an inert gas. Some strong hydrates may hold from losing water until reaching a high temperature (eg, $>100^\circ C$). As a general observation, the dehydrate temperature may serve as a useful indication on the physical stability of the hydrate.

A hydrate may undergo stepwise dehydration, forming a lower hydrate first, followed by losing additional water at a higher temperate. This behavior reflects the different groups of water molecules in the crystal lattice, and the group of water molecules with weaker interactions is lost first. This is the case for eprosartan mesylate dehydrate.[32]

A hydrate may dehydrate and form a crystalline anhydrate directly, which melts at the anhydrate melting point. However, it has also been frequently observed that dehydration leads to the formation of an amorphous/disordered phase, which crystallizes and then melts at a higher temperature. The crystallization and dehydration events may overlap, and the peaks

may not be well developed. In some cases, the first crystallized phase is a metastable form, which converts to a more stable form and melts.

It should be noted that the humidity in the purging gas in the furnace is essential zero so that dehydration in DSC is irreversible in nature. Therefore, it is not surprising that dehydration behavior may change with the sample pan configuration. For example, dehydration is generally delayed to a higher temperature when a sample is prepared in hermetically sealed pans. The dehydration temperature may slightly increase when samples are packed in sealed pans with pinholes than in an open pan where the water vapor may escape freely.

If the RH in the dehydration environment can be controlled, then the critical RH information may be obtained.[33] The idea is to study the dehydration temperature at different relative humidity levels, which allows the critical RH to be extrapolated to the temperature of interest (usually ambient) based on the Van't Hoff equation where the effect of temperature on chemical equilibrium is delineated.

3.4.1.2.4 Characterization of amorphous phases

An amorphous phase has higher apparent solubility, and it has been exploited to improve the bioavailability of many water-insoluble drugs. However, it is thermodynamically unstable and always bears the risk of recrystallization during manufacturing, storage, or during in vivo dissolution. When recrystallization occurs, the product performance may be sacrificed. Therefore, the physical stability of amorphous pharmaceuticals has been a significant focus over the last 20–30 years.[34]

DSC has been used extensively to study the glass transition, molecular mobility, thermodynamics, enthalpy relaxation, and recrystallization of amorphous pharmaceuticals.[35–44]

Glass transition is a second-order phase transition, where the heat capacity shows a step change. Glass transition is also a kinetic phenomenon, reflecting the molecular mobility of the amorphous phase. Therefore, the value of the glass transition temperature, T_g, is dependent on the heat/cooling rate because it reflects the temperate time it takes the molecules to move approximately the same as the experimental time scale. A higher heating/cooling rate results in a higher T_g value.

The difference in the thermodynamic quantities between the amorphous phase and the crystalline counterpart is called the configurational thermodynamic quantity (see Section 2.7). These configurational parameters can be readily derived based on data on heat capacity and melting, and DSC is the instrument of choice for these measurements. TMDSC is particularly useful and convenient for measuring the thermodynamic heat capacity of both amorphous and crystalline materials.[41,42]

Extrapolation of the configurational thermodynamic quantities from supercooled liquid states to lower temperatures results in a negative configurational quantity at temperature (T_K) above the absolute zero. This is the Kauzmann paradox where the metastable amorphous phase is more stable than the crystalline phase. Of course, this would never happen in reality. However, the Kauzmann temperature T_K represents the zero mobility temperature of an ideal glass.

As indicated in chapter "Crystalline and Amorphous Solids," molecular mobility in the supercooled liquid phase generally follows the Vogel–Tammann–Fulcher (VTF) equation:

$$\tau = \tau_0 \exp\left(\frac{DT_K}{T - T_K}\right)$$

where τ is the relaxation time (reciprocal of mobility), and D is a parameter signifying the temperature dependence of the relaxation time or strength of the glass-former (classified as strong or fragile). Molecular mobility in strong glasses typically exhibits Arrhenius-like temperature dependence, while the Arrhenius plot for a fragile glass is significantly nonlinear. The strength parameter may be derived from the apparent activation energy of the glass transition, ΔH, which again is conveniently determined by measuring the temperature dependence of the glass transition temperature using a DSC:

$$\Delta H = -R\frac{d\ln \beta}{d(1/T_g)}$$

$$D = \frac{\Delta H(1 - T_K/T_g)^2}{RT_K}$$

The fragility parameter, m, is related to D and can be determined similarly:

$$m = \frac{\Delta H}{2.303RT_g}$$

The determining of the configurational heat capacity and configurational entropy allows the molecular mobility of amorphous phase[41,42] to be calculated using the Adam–Gibbs model:

$$\tau = \tau_0 \exp\left(\frac{C}{TS_C}\right)$$

where S_c is the configurational entropy, and C is a material-dependent constant. Findings from thermodynamic and crystallization studies have revealed that configurational entropy plays an important role in determining the crystallization tendency of an amorphous phase.[42]

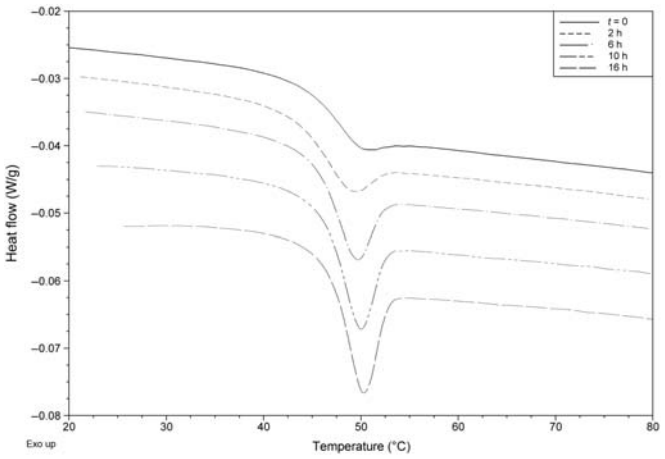


FIGURE 3.14 Enthalpy relaxation of ritonavir glass annealed at 35°C.

Molecular motion at temperatures below T_g, takes longer time, and the resulted enthalpy loss along the process is recovered during reheating in the DSC (Fig. 3.14 as an example). Enthalpy relaxation data has been used to derive the molecular mobility based on the KWW model.[35] We have used enthalpy relaxation to characterize the molecular mobility using the Adam–Gibbs model and the configurational thermodynamic quantities in fresh and annealed glasses. The importance of configurational entropy in glass stability was also confirmed.[44] The configurational entropy has been used to classify the physical stability of amorphous phases in conjunction with the molecular mobility.

Crystallization may be observed above T_g when molecular mobility is large enough. The relative position of the crystallization peak between T_g and T_m can be related to the ease of crystallization tendency.[41] DSC can be used to follow the recrystallization events and extract the kinetic information of the process.[37,43]

3.4.2 Thermogravimetric analysis

As the name suggests, thermogravimetric analysis (TGA) monitors the weight of a sample along with a temperature program. Similarly to DSC, the furnace is generally heated under the purge of an inert gas. TGA is primarily used to monitor thermal decomposition of compounds and is frequently combined with DSC to provide information about the nature of a thermal event, such as dehydration/desolvation, desorption, or melt decomposition.

3.4.3 Microcalorimetry

Isothermal microcalorimetry (IMC) is a highly sensitive calorimetry that is used to study thermal processes at a set temperature in the range of 5–95°C. IMC instrumentation can detect signals in the order of 10^{-6}°C in ΔT, far exceeding the sensitivity in conventional DSC (on the order of 10^4 times increased sensitivity). Thermal reactions that lead to physical and/or chemical stability may be studied at higher temperatures where the processes proceed at significant kinetic rates. However, pharmaceutical products are in general concerned about the stability during normal and accelerated storage conditions (eg, 5–40°C). Extrapolation of kinetic information to temperature and other conditions that are of interest to pharmaceuticals may be challenging and not reliable due to potential changes in reaction pathways or their relative contributions. The introduction of IMC in the 1980s (eg, thermal activity monitor, TAM) has made it possible to monitor heat processes and deduce stability information of pharmaceuticals at temperatures of interest directly.

The heat flow from an IMC experiment can be represented as:

$$\frac{dQ}{dt} = \Delta H \cdot C_0 \cdot \frac{d\alpha}{dt}$$

where ΔH is the enthalpy of the process under investigation, C_0 is the initial reactant concentration, α is the extent of conversion, ranging from 0 to 1, and $d\alpha/dt$ is the rate of the reaction.

IMC has been used to study most typical physical chemical processes of pharmaceuticals, including crystallization, polymorphic transformation, dehydration/hydration, and chemical reaction such as drug decomposition and excipient compatibility in formulation design.[22,45–48] The volume of the IMC cells allows the humidity of the headspace to be controlled through a humidity-stat such as a salt solution. Therefore, it can be useful to study the hygroscopicity, dehydration/hydration of drugs, and other pharmaceutical components.

One drawback of IMC is that it is incapable of temperature scanning. To that end, micro DSC is capable of performing both isothermal and nonisothermal calorimetric studies. Because of the increased sample size and the sensitivity requirements, the scanning rate is generally small (up to about 1°C/minutes), and the temperature ranges are generally limited (from about −40°C to anywhere between 100°C and 200°C). However, to the plus, the instrument allows for a very slow scanning rate such as 0.001°C/minutes to be controlled. Micro DSC has been used to study the phase transition of biological samples, hydration/dehydration of hydrates, and crystallization of amorphous materials under controlled RH, as well as drug degradations.

3.5 VIBRATIONAL SPECTROSCOPY

Spectroscopy is a way to figure out molecular structure, and it primarily uses electromagnetic

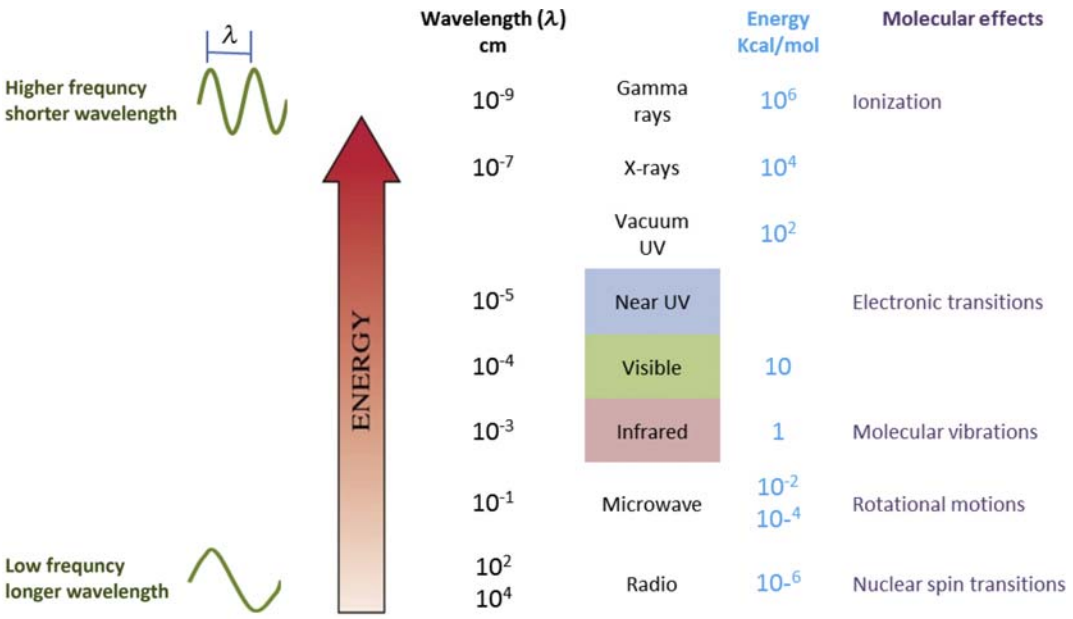


FIGURE 3.15 Regions of electromagnetic spectrum.

radiation. Fig. 3.15 highlights the major regions of the electromagnetic spectrum. Note the larger the wave number, the higher the energy. It is important to note the relationship between wave numbers and wavelength (λ):

$$\text{Wavenumbers} = \frac{1}{\text{Wave length } (\lambda)}$$

Visible light is used for determining the atomic structures of H and He, and both visible and UV light is used to study the electronic structure of atoms and molecules. Infrared light is in the region of energy where there is bond vibration, and thus it can provide information regarding functional groups and some knowledge of the structure of the molecule. The radio spectrum region provides information regarding quantum transitions; thus it provides the most definitive structural information.

Spectroscopic techniques primarily used for solid-state characterization in pharmaceutics are: infrared, Raman, and solid-state nuclear magnetic resonance (SSNMR). Therefore, each of the techniques will be discussed in some detail, including instrumental fundamentals, and then some applications will be used to illustrate the value of the method for solid-state characterization.

3.5.1 IR and Raman spectroscopy

3.5.1.1 IR spectroscopy

When exposed to IR radiation, *absorption* of electromagnetic radiation by molecules occurs depending on the functional groups leading to bond vibration (stretching, contracting, bending, wiggling, and twisting). It is possible to measure how much energy is absorbed at each particular wavelength and make inferences about the types of bonds that are present in the molecule. IR spectrum identifies functional groups by looking at the transitions between quantized vibration states, and it involves the change in dipole moment of the molecule.

TABLE 3.1 Regions of IR Spectrum

Region	Wavenumber (cm^{-1})
Near IR	4000–14,000
Mid IR	500–4000
Far IR	5–500

Infrared region of the electromagnetic spectra can be divided into three main regions, and the mid-IR region is the most useful (Table 3.1).

As already alluded to, there are several types of molecular vibrations. For a molecule to be IR active, it must possess a dipole moment (a permanent dipole). Dipole moment is a vector that is equal in magnitude to the charge multiplied by the distance between the changes and direction toward the net positive charge. Homonuclear diatomic molecules such hydrogen, oxygen, and nitrogen undergo symmetrical stretching and are not IR active, since there is no resultant dipole moment. As an example, Fig. 3.16 depicts stretching vibrations of an oxygen molecule, and it is evident that none of the three

FIGURE 3.16 Symmetrical stretching of oxygen where bonds are represented as springs.

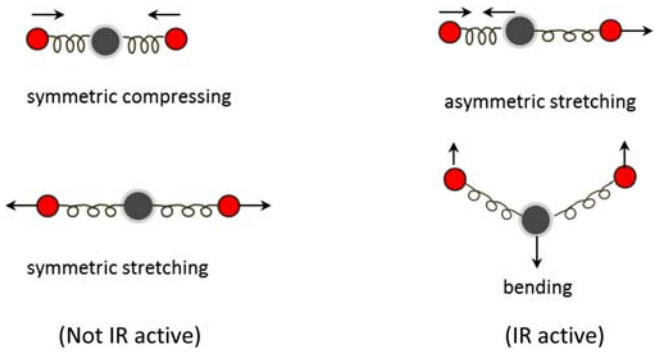


FIGURE 3.17 Carbon dioxide has some vibrations that have a resultant dipole.

configurations would result in a change in dipole; therefore, oxygen is inactive to IR.

On the other hand, molecular vibrations that lead to a resultant dipole are IR active, as illustrated in Fig. 3.17 using carbon dioxide. For the linear geometry with symmetric stretching, the dipole moment is zero, making these configurations IR inactive. However, a carbon dioxide molecule can assume additional configurations: asymmetric stretching and bending that result in positive dipole, and these configurations are IR active.

Other modes of vibration that molecules may exhibit include rocking, wagging, and twisting. The energy difference between the ground state and the higher energy is utilized in generating the IR spectrum. The lower energy state absorbs a photon of energy, that is, a specific wavelength of the electromagnetic band, to go to the higher energy state, but it can give up a photon of energy to return to the ground state. This absorption of energy creates the peak in the IR spectrum. In the case of organic molecules, there are many bonds that have vibrations where photons fall within the IR band of the electromagnetic spectrum. When monochromatic electromagnetic radiation (radiation of a specific wavelength) is passed through a sample, if the wavelength of light matches the energy of bond vibration, then energy will be absorbed, leading to more of that specific vibration that leads to resonance. As the sample is scanned across the IR band, the energy will come into resonance with specific vibrations, and absorption will occur. This absorbed energy produces the peak. The intensity of the peak depends on both the change in dipole moment and the number of specific bonds that are involved.

The simplest form of an IR spectrophotometer involves a source of IR radiation, a sample holder, grating, and a detector. However, there are instrument-related noises that need to be minimized to improve signal; therefore, the double-beam setup described in Fig. 3.18 is common.

In Fig. 3.18, the IR beam is split, with half going through the sample and other half going through the reference. A set of mirrors direct the beam coming from the sample/reference to a rotating chopper. The chopper consists of a transparent end and a mirrored end. The transparent end allows the light from the sample to go through the slit, and the mirrored end reflects the reference light toward the slit. In other words, the alternating light (once sample, then reference) goes through the slit to the grating. The grating then selects the monochromatic light that passes on to the mirror, alternating between the sample and the reference. This light finally reaches the detector, and the difference between the reference and sample light arising from absorption of light by the sample is translated to the IR spectra. It is worth noting that Fig. 3.18 depicts a dispersive IR spectrophotometer, which is rarely or never used. Today the spectrophotometers are FTIR spectroscopy because they offer better signal-to-noise ratio, better frequency accuracy, and faster acquisition and analysis time.

The key to FTIR spectroscopy is the interferometer, and a schematic is provided in Fig. 3.19, which will be used to discuss the functioning of the equipment.

Let's start by just looking at the sample. The light (assume monochromatic for now) from the source goes

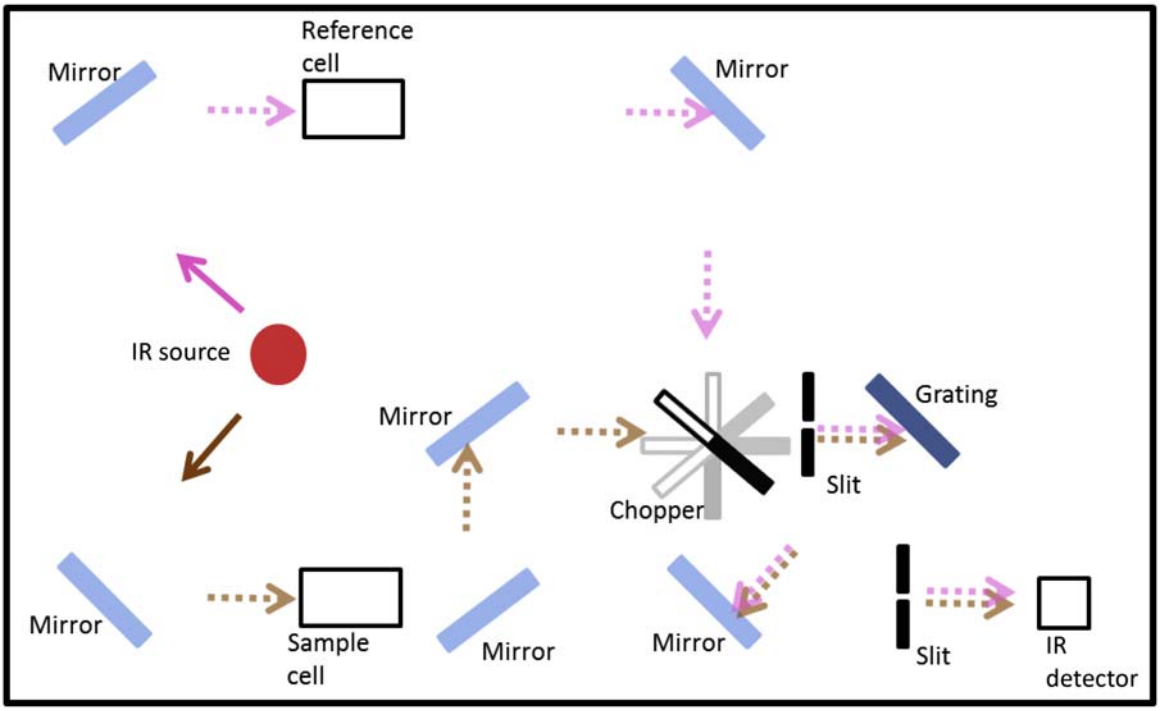


FIGURE 3.18 Double-beam dispersive IR spectrophometry.

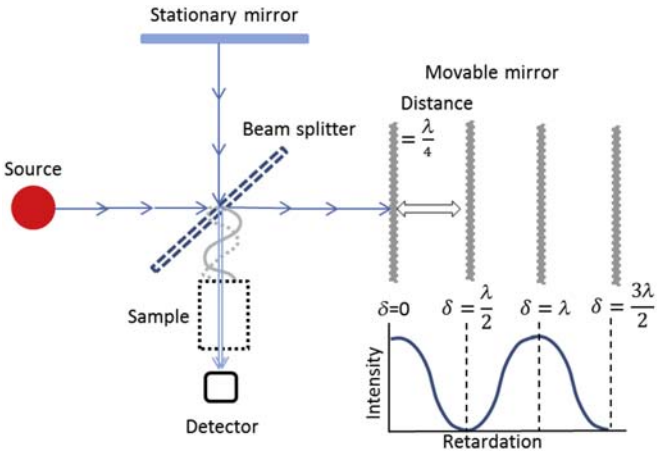


FIGURE 3.19 Schematic of an interferometer.

to a beam splitter, which reflects some of the light to the stationary mirror and transmits the rest (say 50/50 for now) to the movable mirror. The rays are then reflected back to the beam splitter where again half of each ray is reflected and other half transmitted. In other words, the ray that came back from the stationary mirror will be split, with half going back to the source and half going through the sample to the detector. The same happens to the ray returning from the movable mirror. The two optical paths, beam splitter to stationary mirror and beam splitter to movable mirror, are not usually equal; therefore, the light going to the samples will not always be in phase. This means there will be interference.

Two times the difference in the path between the stationary and movable mirror is called retardation (δ). If δ is an integer value of wavelength λ, then the interference is constructive, and if δ is an integer value of $\lambda/2$, then the interference is destructive interference. In between those two values, the interference is variable. The output graph, intensity versus retardation, is the interferogram. The interferogram from the reference and the sample are Fourier transformed to produce the spectrum. Note in reality the light from the source is a broad-spectrum IR source, and the beam splitter has an efficiency associated with its ability to split the beam, and this efficiency is an instrument constant.

3.5.1.2 *Raman spectroscopy*

Complementary to infrared spectroscopy is Raman spectroscopy. Raman is a *scattering* phenomenon that occurs when molecules interact with an electromagnetic field. Light that is scattered from a molecule is primarily elastically scattered; that is the incident and the scattered photons have the same energy, and this is called Rayleigh scatter. Rayleigh scattering is not useful for analysis, and most instruments filter out this elastic scattering. However, an exceedingly small fraction of photons maybe scattered inelastically, resulting in either a net gain or loss of energy of the scattered photon.

TABLE 3.2 IR versus Raman Spectroscopy

IR	Raman
Results from absorption of light by vibrating molecule	Results from scattering of light by vibrating molecule
Vibration is IR active if there is a change in dipole moment during vibration	The vibration is Raman active if it causes a change in polarizability
Molecule possess a permanent dipole	Molecule does not need to possess a permanent dipole
Water cannot be used as solvent	Water can be used as solvent
Sample preparation requires care. Gaseous samples can rarely be used	Sample preparation is simple

The Raman process can be described as a molecule in an initial vibrational state that is excited to a higher energy state and then returns to a vibrational state (higher or lower than initial) with simultaneous scattering of a new photon from this state. The difference in energy between the incident and scattered photon is equal to the energy difference between the initial and final vibrational states of the molecule. A loss in photon energy (ie, when it returns to a lower energy state) is termed Stokes–Raman scattering, and a gain in photon energy or when it returns to a higher energy state is termed anti-Stokes–Raman scattering. Not all vibrational transitions can be accessed by Raman scattering. Raman-active transitions are those that are associated with a change in polarizability of the molecule. Polarizability is the ease with which the electron cloud can be distorted by an applied field. In other words, Raman activity is associated with induced dipole. A Raman spectrum is a plot of scattered light intensity versus Raman shift reported in wavenumbers (cm^{-1}).

As stated before, IR and Raman are complementary techniques, with some similarities and differences. Table 3.2 summarizes the differences, and Fig. 3.20 illustrates the advantages of using these two techniques in conjunction in order get information regarding the functional groups.

The use of these two techniques is illustrated using carbon tetrachloride. CCl4 has a tetrahedral configuration with the chlorine atoms being more electronegative than the carbon atom. The scissoring, rocking, and symmetric stretching vibrations do not result in molecular dipole moment, and, therefore, they are IR inactive, but polarizability of the electron cloud results in a Raman shift. Only the asymmetric stretching vibration produces an induced dipole; therefore, this vibration is IR active.

In the field of solid-state pharmaceutics, these techniques are used not only for elucidating functional groups but to find far greater use in a variety of applications that span research that focuses on elucidating molecular packing associated with various crystal forms[49–52]; uses that lead to material identification[53]; formulation performance, which may be solid properties,[54–57] release, or stability[58]; manufacturing tool for endpoint determination[59,60] or for looking at phase transformations[61,62]; and monitoring counterfeit products,[63–65] to name a few. Since these techniques are so widely used in pharmaceutical manufacturing, there are many reviews that discuss pharmaceutical applications of IR and Raman spectroscopy.[66,67] It is evident that that method is a valuable tool for pharmaceutical analysis, and if one were to mention any disadvantage, it is less sensitive compared to PXRD in distinguishing between crystal forms or between crystalline and amorphous.

3.5.2 SSNMR spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy is a technique that exploits the magnetic properties of certain nuclei to study physical, chemical, and biological properties. It involves the absorption of electromagnetic radiation in the radiofrequency (rf) region. The basic principles covered utilize minimum mathematics and no quantum mechanics; therefore, the reader should refer to specialized textbooks for details.[68–70] Briefly, atoms that have an odd number of protons, neutrons, or both can behave like magnets (Fig. 3.21a), for example, ^{1}H, ^{13}C, ^{15}N, ^{17}O, and ^{19}F. The strength of their magnetic field is measured by a vector called magnetic moment, which is related to a fundamental nuclear property called the nuclear spin (I). The rules for determining the net spin of a nucleus are as follows:

1. If the number of neutrons and the number of protons are both even, then there is no spin, and it is not NMR active.
2. If the number of neutrons plus protons is odd, then the spin is half-integer (ie, 1/2, 3/2, 5/2).
3. If the number of neutrons and protons are both odd, then the spin is an integer (ie, 1, 2, 3).

A nucleus of spin I has $2I+1$ orientations. Therefore, ^{1}H atom with $I=1/2$ has two possible orientations ($I=1/2$, $I=-1/2$). In the absence of an external magnetic field, the magnetic moments of the

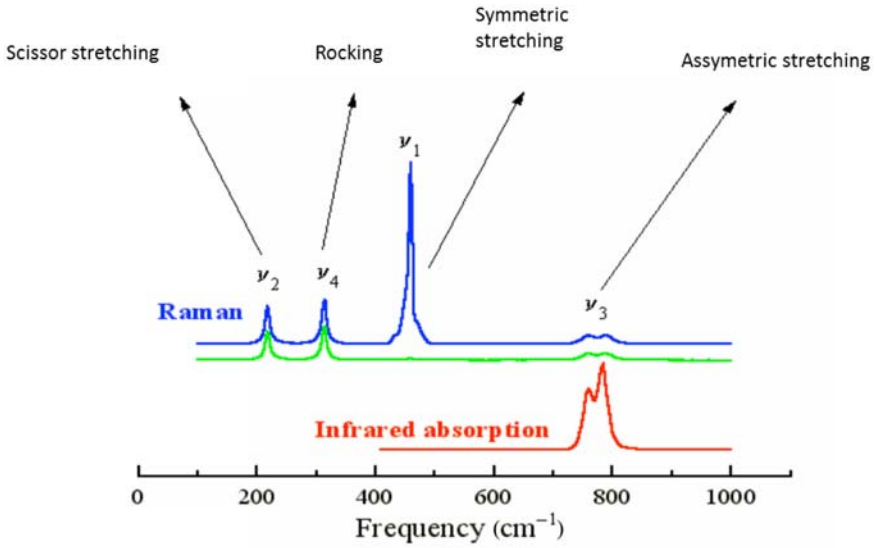


FIGURE 3.20 Raman and IR spectrum for carbon tetrachloride. Taken from http://www.ccmr.cornell.edu/igert/modular/docs/Raman_Scattering.pdf

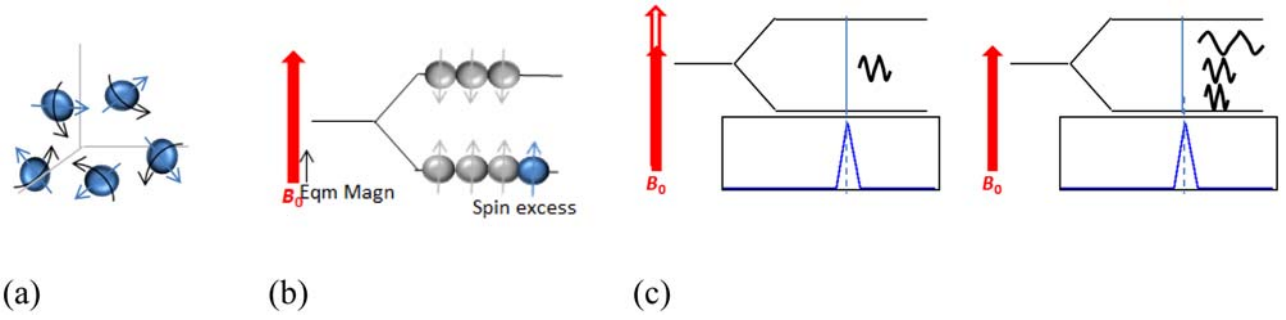


FIGURE 3.21 (a) Nucleus (ex. ^1H, $I = 1/2$) with an angular momentum randomly oriented. (b) Splitting of nuclear spin under applied magnetic field B_0, producing a small magnetization. (c) Two techniques for continuous wave NMR: The first increases the magnetic field while a constant frequency is on, while the second does a frequency sweep at constant B_0. Absorption occurs at Larmor frequency.

nuclei are random, or there is isoenergic distribution of nuclear spin. The distribution is described as a Boltzmann distribution.

When a magnetic field (B_0) is applied (Fig. 3.21b), the energy levels split (Zeeman splitting), and the nucleus (the "magnets") will line up with the external magnetic field. There is usually a slight preponderance of excess spin at the lower energy state that results in equilibrium magnetization (Fig. 3.21b), which aligns with the external magnetic field. This excess is detected as a signal, but it is a weak signal. If, on the other hand, the sample size is large, there could be an abundance of this spin excess nucleus. When an oscillating energy with a frequency equal to the difference in energy states is applied while B_0 is increased, absorption occurs. Alternatively B_0 can be held constant, and a frequency sweep can be applied, and when there is a match in frequency (or resonance occurs), there is absorption. This absorption results in NMR signal (Fig. 3.21c). However, these types of experiments, even with a large sample size, give weak signals.

Two important concepts, relaxation time and precession, need to be discussed. At equilibrium, the net magnetization vector lies along the direction of the applied magnetic field B_0 and is called the equilibrium magnetization (Figs. 3.21b or 3.22a). It is possible to change the net magnetization by exposing the nuclear spin system to the energy of a frequency equal to the energy difference between the spin states. If enough energy is put into the system, it is possible to saturate the spin system and make this magnetization zero. Then when the energy source is removed, the magnetization vector will return to the original equilibrium state, and the time constant of this transition is called the spin lattice relaxation time, T_1.

Since the particles are spinning in the presence of this applied magnetic field (B_0), the net magnetization vector will not only be in the z-direction, but it will rotate about the z-axis, a phenomenon called precession. This precession is evident only when the axis of the nucleus is not exactly aligned with the magnetic field. The precision frequency (ω) or Larmor frequency

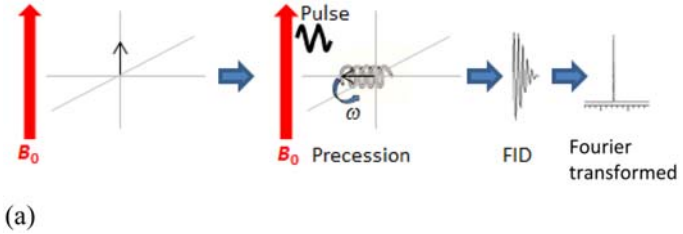


FIGURE 3.22 (a) NMR spectroscopic configuration. (b) Shielding due to secondary field generated by electron cloud.

is independent of the orientation of the nucleus to the magnetic field but dependent on B_0 and the gyromagnetic ratio (γ) of the nucleus. The γ is a characteristic of the nucleus. Note:

$$\omega = \gamma B_0$$

This equation relates precession frequency (ω) or Larmor frequency to the applied magnetic field and the gyromagnetic ratio of the nucleus and is the most important equation for NMR spectroscopy. When the oscillating energy orthogonal to B_0 is introduced, it tips the nuclei out of its equilibrium position, causing a transition between the two energy levels of the spin. When the energy source is removed, the spin relaxes to the equilibrium state with a spin−spin time constant T_2. As spin relaxation occurs, the net magnetization reaches equilibrium. It should be noted that Fig. 3.22a depicts a 90° pulse technique, which has the problem that the rf field must be turned off before signal can be collected. There are other techniques used to overcome this problem that are beyond the present scope.

In NMR, the energy source is a probe coil that produces the rf signal (transmitter), and as the nucleus returns to the equilibrium state, the signal that decays with time is in the form of an oscillating voltage and is picked up by the same coil, which then acts as a receiver. In the most commonly used NMR spectroscopic configuration, the signal is in time domain and is called the free induction decay. It is Fourier transformed into frequency domain to provide the NMR signal.

Once the spectrum is obtained in the frequency domain, there are two types of information that can be gleaned from it: the amount of nuclei generating the spin, which can be achieved by integrating the peak, and the chemical environment of the nucleus generating the signal, which determines the position of the peak in the NMR spectrum. The chemical environment where there are other electrons are present can lead to shielding of the nuclei, which requires less energy to bring it into resonance, and conversely deshielded nuclei require more energy, and the peaks appear downstream. It is worth noting that since the shielding tensor is dependent on the applied magnetic field, a chemical shift needs to be calibrated with a reference.

In solid-state NMR spectroscopy, molecular motions are restricted to local vibrations and rotations; therefore, dipolar interactions are the primary interactions. In the absence of rapid molecular tumbling that leads to average conformation in a solution, in powder samples, the NMR spectrum consists of a superposition of numerous lines that originate from a given nucleus. This leads to excessive peak broadening. Such broad peaks are not useful. Therefore, ^{1}H NMR is not used for obtaining chemical shift information when the sample is a powder. Since the chemical shift anisotropy leads to broad peaks from overlapping patterns and interplay between different broadening mechanisms, the elimination of anisotropic interactions in high-resolution SSNMR is achieved by spinning the sample around an axis inclined at 54.74°, known as magic angle. Since dipolar interactions are weakened by fast rotation, rotor spin rates are selected such that they are larger than the anisotropic interactions. Magic angle spinning (MAS) is combined with decoupling to improve spectral resolution. Decoupling consists of an application of short pulses that prevent line splitting of the less abandoned nuclei. SSNMR sensitivity for a low-sensitive nucleus (eg, ^{13}C) can additionally be increased by using a cross-polarization (CP) technique. It is based on excitation of the sparse spin system through the abundant one. In the case of ^{1}H-^{13}C CP, if the Larmor frequencies ω_H and ω_C are made equal by adjusting the oscillating field amplitude (satisfying the Hartmann−Hahn condition $\gamma_H B_H = \gamma_C B_C$), then resonance exchange of energy between the two types of nuclei can easily take place. In fact, this condition can be considered as thermal exchange, and the time taken to cool down ^{13}C is associated with the warming of ^{1}H and is called contact time (CT). A 90° pulse is applied, and then the two spin systems are put in thermal contact through a pulse train having CT duration that allows Hartmann−Hahn matching. Finally the ^{1}H nucleus is decoupled, and acquisition of ^{13}C signal is performed. SSNMR of adamantine in Fig. 3.23 demonstrates the advantages of CP MAS NMR.

High-resolution SSNMR can be used to extract both structural and dynamic information, but such experimentation requires significant expertise, expensive instrumentation, large sample size, and is time consuming. However, this powerful, nondestructive

but theoretically complex tool has been used for qualitatively and quantitatively analyzing crystalline and amorphous pharmaceutical samples and for characterizing both active ingredients as well as formulations. SSNMR is widely used in pharmaceutical applications, and there are several reviews that go into further depth with regard to theory and application.[72–74] A few applications are highlighted here. In conjunction with single crystal structural information, SSNMR of erythromycin and clarithromycin was analyzed to gain understanding for the propensity of these molecules to form hydrates and solvates.[75] The stability of an amorphous statin was investigated by solid SSNMR, and a strong dynamic heterogeneity was attributed to the stability of simvastatin.[76] Investigations that involved formulations have used SSNMR for studying polymorphism in dosage forms when the active ingredient is present in low doses.[77] In the development of amorphous solid dispersion, the solubility of nifedipine in PVP was investigated by SSNMR.[78]

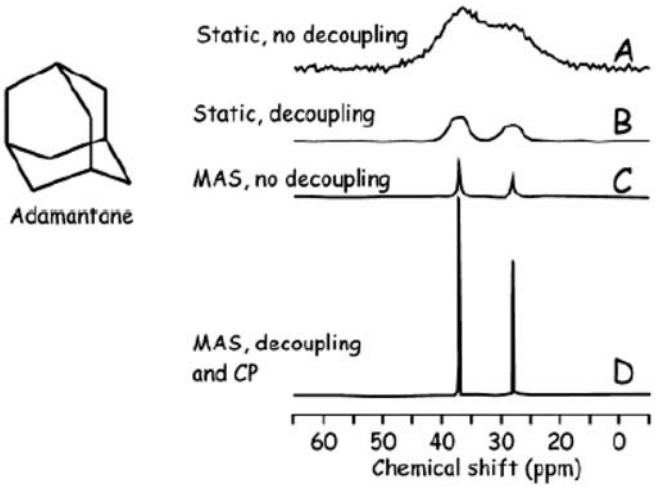


FIGURE 3.23 SSNMR spectrum of adamantine.[71]

3.6 MOISTURE SORPTION

Vapor sorption is a sensitive technique that is extensively used for characterizing API's, excipients and formulations. It can be destructive if there is extensive loss or gain of weight during testing. However, the method is simple, requires a tenth of a sample as compared to PXRD, and is almost 10-fold more sensitive in detecting the presence of a mixture of phases, making it an attractive tool for routine utilization, but it has applications other than phase purity evaluation.

The concept and the equipment design are amazingly simple. The method relies on the concept that water sorption is a function of water activity (a_w). Water activity (a_w) is defined by the ratio of vapor pressure of water (p_w) divided by vapor pressure of pure water (p_o) at a given temperature, and 100 times that ratio is relative humidity (%*RH*):

$$a_w * 100 = \frac{p_w}{p_o} * 100 = \%RH$$

Instrumentation involves a temperature-controlled chamber, a few flow meters, and a microbalance to which are attached a reference and a sample pan (Fig. 3.24). In Fig. 3.24a, the chamber is kept at constant temperature. Dried nitrogen (N_2) flows into the chamber and is split into two streams. One stream passes over a solvent, which is usually water, producing a stream of wet N_2 at 100% RH. Then the wet and dry N_2 are mixed at a given ratio controlled by the

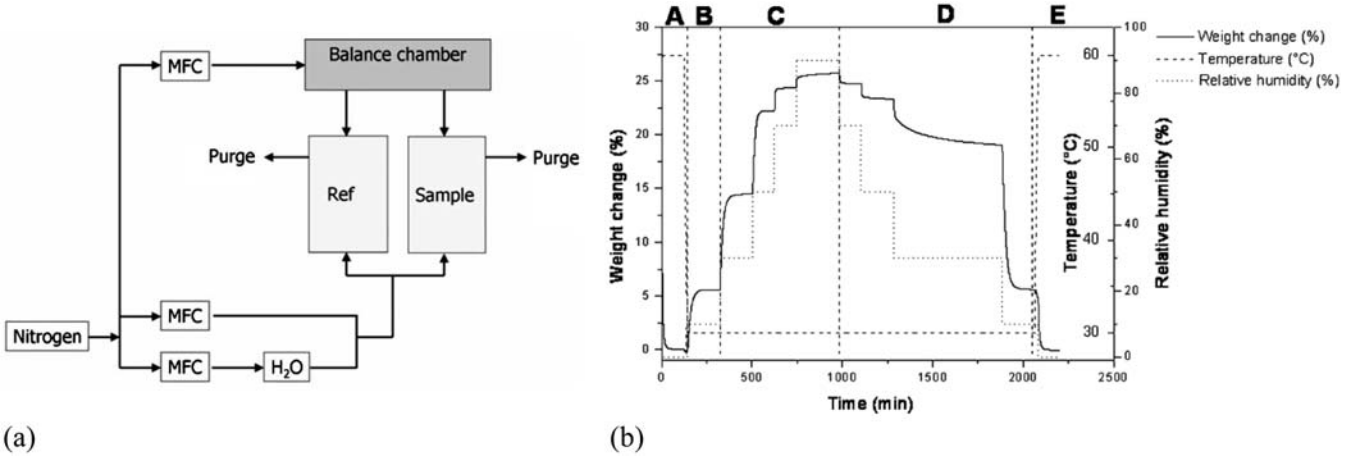


FIGURE 3.24 (a) Experimental setup for dynamic vapor sorption gravimetric analysis. (b) An example: isothermal measurement of silica.[79]

respective mass flow controllers to obtain a N_2 of a given relative humidity. This N_2 of a defined humidity flows through the reference and sample chambers.

The sample is first dried and then exposed to changing humidity values, and the weight change is recorded as a function of humidity. Usually the humidity is increased from 0% to 90% RH in steps of 10% or 20% RH between equilibration times. Finally the humidity is decreased in steps of 10% or 20% RH. Following Fig. 3.24b, the sample is first dried at 0% RH when there is weight loss, as depicted by the solid line in segment A. In segment B, the dotted line indicates that the RH was ramped up by 20%, and the solid line indicates that the initial sample rapidly gains weight, and then there is no appreciable weight change (ie, the sample reaches an equilibrium at 20% RH). Between segments B and C, there are four more steps of humidity increase (dotted line), and at each step, the sample, after initially gaining weight, stabilizes to the humidity conditions. In segment D, the humdity is decreased in steps, and at each step, the sample loses weight. The sample equilibrates at each step before the RH is changed.

The abscissa in Fig. 3.24b is in time, and it is evident that the stabilization at a given RH is dependent on whether appreciable weight gain or loss occurs at that RH. The goal is to ensure equilibrium weight at each RH; thus the greater the weight change, the longer the time for the step. It is common to plot this weight change data as a function of %RH. These plots provide valuable information, and some examples have been provided in Fig. 3.25, mainly to demonstrate that it is both the extent of moisture uptake/loss and the nature of the plots that provide valuable information. It is also common to present this data without including segment A (ie, the weight change of the sample is normalized to dry basis). Finally only two examples will be used to demonstrate the types of information that can be obtained from moisture uptake profiles.

Fig. 3.25a is an example that depicts the moisture uptake profile of amorphous and crystalline solids. At any given RH, the moisture uptake by the amorphous is greater than the crystalline counterpart. The adsorption (upward solid arrow) and desorption (downward solid arrow) curves exhibit a hysteresis or a shift to higher water content at a given %RH during desorption. There are a few theories associated with the existence of this hysteresis for adsorption-desorption curves, and they are beyond the scope of this present text. However, it is important to note that most materials that undergo adsorption-desorption without any phase change may exhibit such hysteresis.

When phase change occurs, the moisture uptake profiles become complicated. In the example in Fig. 3.25b, the amorphous rapidly takes up 10% (w/w) water up to 80% RH. At 80% RH, there is 5% (w/w) weight loss, and even when the humidity is increased to 90% RH, the weight gain of the sample remains at ~5% (w/w). During desorption the sample remains almost unchanged from 90% to 30% RH. Between 30% and 10% RH, the sample has ~3% (w/w) water content. This type of sample behavior is typical for amorphous samples that undergo crystallization. Initially, the amorphous rapidly absorbs water, similar to what was seen in Fig. 3.25a, but since water is a potent plasticizer, in this case it promotes crystallization. Whether crystallization occurs or not is a material property. The crystallizing solid loses water, leading to the observed

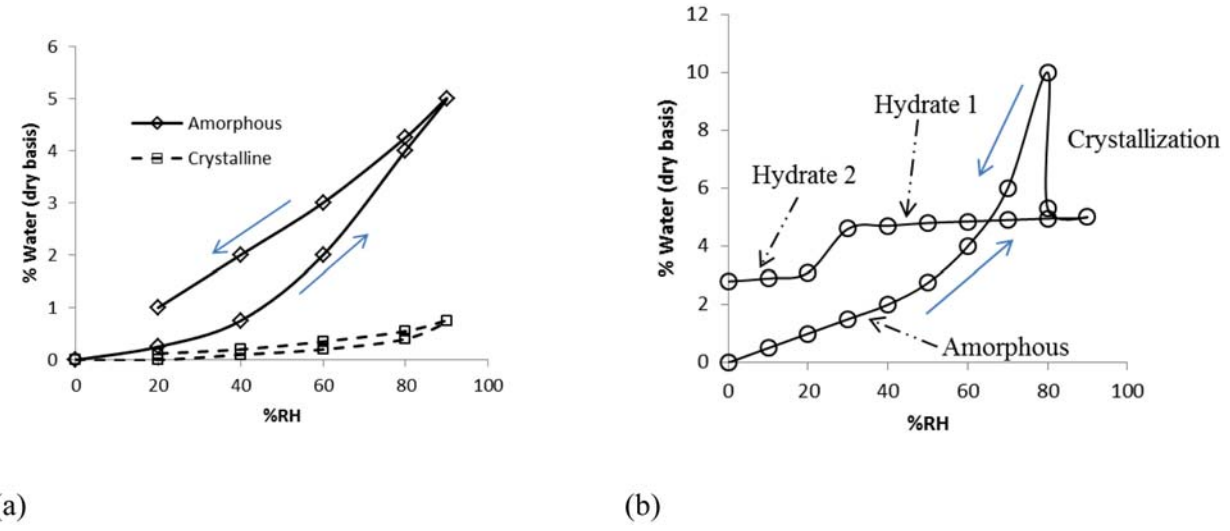


FIGURE 3.25 Examples of dynamic vapor sorption gravimetric analysis of (a) amorphous and crystalline solids with no phase change and (b) amorphous undergoing moisture mediated crystallization.

weight loss. In this example, the solid after weight loss at 80% RH contains 5% (w/w) water; therefore, it can be assumed that a hydrate was formed. In other words, the water in the sample at this stage is no longer sorbed water but has become a part of the lattice. If it were an anhydrous phase, it would have lost almost all of the water it had originally gained. Finally, there is another step change that occurs at 30% RH, indicating the formation of a lower hydrate containing ~3% (w/w) water. These moisture sorption profiles give an indication of changes as a function of water uptake/loss, but these changes need to be confirmed by orthogonal techniques such as PXRD. Fig. 3.25 not only provides information regarding the existence of different solid phases (amorphous, hydrate 1, hydrate 2), it also provides information regarding the hygroscopicity of each solid phase. Again it is evident that the moisture uptake rate by the amorphous is greater than either of the crystalline solids.

The moisture uptake technique described above is very widely used in the pharmaceutical industry, and some examples have been cited, but there are a few other techniques that are widely used in drug development that are worth mentioning. Since electrolytes depress the vapor pressure of water at a given temperature, a saturated salt solution will have a defined RH. Therefore, a cheap gravimetric method capitalizing on the above concept (ie, water uptake by a solid) is a function of a_w of the environment, which utilizes desiccators and salt solutions for exposing the sample to various humidity conditions.

The other two techniques that are commonly used when the total water content of a sample needs to be determined are: loss on drying (LOD) and Karl Fischer titration. Both these are common compendial methods. In the case of LOD determination, a given weight of the sample is heated to a defined temperature, and the weight loss is recorded. Therefore, this LOD method is also a gravimetric technique, but it cannot provide indication of the presence or absence of hydrates or solvates. Sometime orthogonal techniques such as PXRD before and after LOD determination are used to investigate phase change. It is worth noting that this LOD method does not provide information on whether the weight loss is associated with water, solvent, or decomposition.

In Karl Fischer titration, water in a sample can be determined by either volumetric or coulometric titration. Volumetric titration is used when the sample contains a relatively large amount of water (1−100 mg being titrated), while the coulometric method is used for titrating microgram quantities of water. Briefly, the titration is performed in an alcoholic solution (methanol) in the presence of a base, sulfur dioxide, and a known quantity of iodine. One mole of iodine consumes a mole of water. As is evident, Karl Fischer titration is more specific to water determination than the LOD method, but it also suffers from the disadvantage that it cannot differentiate between surface water and water present as a part of the crystal structure.

In the pharmaceutical industry, information regarding moisture content and moisture uptake is used for several applications, and here are a few examples: solid form selection (ie, polymorph and salts[80–83]), characterizing amorphous formulations,[84,85] detecting process-induced phase transformations,[86] and evaluating packaging configurations.[87]

3.7 HYPHENATED TECHNIQUES

The above techniques are often used in combination, and when they are combined, they provide a multifaceted analytical tool for solid-state characterization. For example, thermal analysis using DSC affords information regarding change; for example, an endotherm occurs for loss of residual solvent, melting, polymorphic transition, desolvation, or decomposition. Although it affords quantitative information regarding heat consumed during the change, it affords no information regarding the type of transition. Hot-stage microscopy involves a heating stage combined with microscopy such that the sample is heated at a given rate on a microscope stage allowing one to observe melting, polymorphic transition, or decomposition. If the experiments are performed with the sample suspended in oil, it is also possible to observe dehydration and even differentiate between residual solvent loss and desolvation. Although hot-stage microscopy does not provide any quantitative data, it allows for precise identification of the type of transition. Another common technique is variable-temperature PXRD, which has been commonly used to study transformation of one crystal form to another as the sample is heated and repeatedly scanned at a predefined rate to generate PXRD pattern. In other cases, a controlled humidity purge has been used in conjunction with PXRD to study the kinetics of moisture-mediated transformation or the kinetics of hydrate formation. Likewise it is possible to combine other techniques such as SSNMR with variable temperature to gain insight. Combined thermal methods such as DSC/TGA have also been commercialized so that both calorimetric and gravimetric information may be obtained in a single scan. When combined with a mass spectrometer (MS) or an FTIR spectrometer, the technique allows the nature of the gaseous reaction products formed in the TGA to be investigated directly and identified directly, even in the presence of multiple volatile degradants. Thus

TGA-MS and TGA-FTIR can be useful in the identification of the starting material and degradation, including kinetics and pathways. Raman microscopy has been used for evaluating uniformity of API in pharmaceutical formulation and especially in the case of amorphous solid dispersion formulations where it is imperative that the API is homogeneously dispersed in the polymer matrix. This technique has also been used to monitor the crystallization or physical stability of amorphous solid dispersions as a function of heat and humidity. These are only a few examples of hyphenated techniques. Advanced instrumentation is available that combines several tools together in order to obtain as much information as possible from a single experiment.

3.8 CONCLUSION

Solid dosage form development often involves the characterization of raw materials going into the manufacturing of the API so that the correct form for development can be selected, characterization of the excipient, intermediates in dosage form manufacturing, and the finished product during release, stability and on-market. Since every imaginable step in development relies on testing such that the material is well understood, there is a plethora of techniques that pharmaceutical scientists utilize. This chapter provides a flavor of some of the common techniques used in support of pharmaceutical development and research. As stated in the introduction, there are a number of other techniques that have been left out. However, the goal is to provide an understanding of some of the common and somewhat complex tools with respect to instrumentation and application.

References

1. Murphy, D.B. and Davidson, M.W. Fundamentals of Light Microscopy and Electronic Imaging, 2nd Ed. Wiley-Blackwell, Hoboken, NJ, 2012.
2. Yu L. Polymorphism in molecular solids: an extraordinary system of red, orange, and yellow crystals. *Acc Chem Res* 2010;**43**(9):1257–66.
3. Greco K, Bogner R. Crystallization of amorphous indomethacin during dissolution: effect of processing and annealing. *Mol Pharm* 2010;**7**(5):1406–18.
4. Chen X, Morris KR, Griesser UJ, Byrn SR, Stowell JG. Reactivity differences of indomethacin solid forms with ammonia gas. *J Am Chem Soc* 2002;**124**(50):15012–19.
5. Wardrop J, Law D, Qiu Y, Engh K, Faitsch L, Ling C. Influence of solid phase and formulation processing on stability of abbott-232 tablet formulations. *J Pharm Sci* 2006;**95**(11):2380–92.
6. Mu L, Feng SS. A novel controlled release formulation for the anticancer drug paclitaxel (Taxol®): PLGA nanoparticles containing vitamin E TPGS. *J Control Release* 2003;**86**(1):33–48.
7. Ward MD. Bulk crystals to surfaces: combining X-ray diffraction and atomic force microscopy to probe the structure and formation of crystal interfaces. *Chem Rev* 2001;**101**(6):1697–726.
8. Lauer ME, Grassmann O, Siam M, Tardio J, Jacob L, Page S, et al. Atomic force microscopy-based screening of drug-excipient miscibility and stability of solid dispersions. *Pharm Res* 2010;**28**(3):572–84.
9. Brittain HG. *Polymorphism in pharmaceutical solids*. Delhi: CBS Publishers & Distributors; 2009.
10. Jenkins RL, Snyder R. *Introduction to X-ray powder diffractometry*. New York, NY: Wiley-Interscience Publication; 1996.
11. Ladd M, Palmer R. *Structure determination by X-ray crystallography*. New York, NY: Springer; 2013.
12. Desiraju GR. Crystal engineering: a holistic view. *Angew Chem Int Ed* 2007;**46**(44):8342–56.
13. Raw AS, Yu LX. Pharmaceutical solid polymorphism in drug development and regulation. *Adv Drug Deliv Rev* 2004;**56**(3):235–6.
14. Stevenson CL, Bennett DB, Lechuga-Ballesteros D. Pharmaceutical liquid crystals: the relevance of partially ordered systems. *J Pharm Sci* 2005;**94**(9):1861–80.
15. Clearfield A, Reibenspies JH, Bhuvanesh N. *Principles and applications of powder diffraction*. New York, NY: John Wiley & Sons; 2009.
16. Anwar J, Barnes P. Kinetics of phase transformations in crystals of drug compounds using time-resolved powder x-ray diffraction. *Phase Trans* 1992;**39**(1–4):3–11.
17. Sundaramurthi P, Suryanarayanan R. Azithromycin hydrates-implications of processing-induced phase transformations. *J Pharm Sci* 2014;**103**(10):3095–106.
18. Uvarov V, Popov I. Development and metrological characterization of quantitative X-ray diffraction phase analysis for the mixtures of clopidogrel bisulphate polymorphs. *J Pharm Biomed Anal* 2008;**46**(4):676–82.
19. Sarode AL, Wang P, Obara S, Worthen DR. Supersaturation, nucleation, and crystal growth during single- and biphasic dissolution of amorphous solid dispersions: polymer effects and implications for oral bioavailability enhancement of poorly water soluble drugs. *Eur J Pharm Biopharm* 2014;**86**(3):351–60.
20. Maurin JK, Pluciński F, Mazurek AP, Fijałek Z. The usefulness of simple X-ray powder diffraction analysis for counterfeit control—the Viagra® example. *J Pharm Biomed Anal* 2007;**43**(4):1514–18.
21. Brown DE. HM. Pharmaceutical applications of thermal methods: the use of DSC and TG in the characterisation of organic and inorganic hydrates. *Thermochim Acta* 1985;521–4.
22. Buckton G, Russell SJ, Beezer AE. Pharmaceutical calorimetry: a selective review. *Thermochimica Acta* 1991;**193**:195–214.
23. Gaisford S, Buckton G. Potential applications of microcalorimetry for the study of physical processes in pharmaceuticals. *Thermochimica Acta* 2001;**380**(2):185–98.
24. Giron D. Thermal analysis and calorimetric methods in the characterisation of polymorphs and solvates. *Thermochimica Acta* 1995;**248**:1–59.
25. Boersma SL. A theory of differential thermal analysis and new methods of measurement and interpretation. *J Am Ceram Soc* 1955;**38**:281–4.
26. Burger AKK. Polymorphism and pseudopolymorphism on nifedipine. *Sci Pharm* 1996;**64**:293–301.
27. Yu L, Stephenson GA, Mitchell CA, Bunnell CA, Snorek SV, Bowyer JJ, et al. Thermochemistry and conformational polymorphism of a hexamorphic crystal system. *J Am Chem Soc* 2000;**122**(4):585–91.
28. Burger A, Ramberger R. On the polymorphism of pharmaceuticals and other molecular crystals. I. *Microchimica Acta* 1979;**72**(3):259–71.

29. Burger A, Ramberger R. On the polymorphism of pharmaceuticals and other molecular crystals. II. *Microchimica Acta* 1979;**72**(3):273–316.
30. Yu L. Inferring thermodynamic stability relationship of polymorphs from melting data. *J Pharm Sci* 1995;**84**(8):966–74.
31. Khankari RK, Grant DJW. Pharmaceutical hydrates. *Thermochimica Acta* 1995;**248**:61–79.
32. Sheng J, Venkatesh GM, Duddu SP, Grant DJW. Dehydration behavior of eprosartan mesylate dihydrate. *J Pharm Sci* 1999; **88**(10):1021–9.
33. Telang C.H.J.; Suryanarayanan, R. Use of humidity controlled DSC (HC-DSC) as a technique to generate rapid physical stability phase diagrams of hydrates. In. NATAS Annu Conf Therm Anal Appl 31st:054/051-054/058; 2003.
34. Hancock BC, Zografi G. Characteristics and significance of the amorphous state in pharmaceutical systems. *J Pharm Sci* 1997;**86**(1):1–12.
35. Hancock BC, Shamblin SL, Zografi G. Molecular mobility of amorphous pharmaceutical solids below their glass transition temperatures. *Pharm Res* 1995;**12**(6):799–806.
36. Andronis V, Zografi G. The molecular mobility of supercooled amorphous indomethacin as a function of temperature and relative humidity. *Pharm Res* 1998;**15**(6):835–42.
37. Aso Y, Yoshioka S, Kojima S. Relationship between the crystallization rate of amorphous nifedipine, phenobarbital, and flopropione, and their molecular mobility as measured by their enthalpy relaxation and ^{1}H NMR relaxation times. *J Pharm Sci* 2000;**89**:408–16.
38. Hancock BC, Christensen K, Shamblin SL. Estimating the critical molecular mobility temperature (TK) of amorphous pharmaceuticals. *Pharm Res* 1998;**15**(11):1649–51.
39. Hancock BC, Dalton CR, Pikal MJ, Shamblin SL. A pragmatic test of a simple calorimetric method for determining the fragility of some amorphous pharmaceutical materials. *Pharm Res* 1998;**15**(5):762–7.
40. Hancock BC, Shamblin SL. Molecular mobility of amorphous pharmaceuticals determined using differential scanning calorimetry. *Thermochimica Acta* 2001;**380**(2):95–107.
41. Shamblin SL, Tang X, Chang L, Hancock BC, Pikal MJ. Characterization of the time scales of molecular motion in pharmaceutically important glasses. *J Phys Chem B* 1999;**103**(20): 4113–21.
42. Zhou D, Zhang GGZ, Law D, Grant DJW, Schmitt EA. Physical stability of amorphous pharmaceuticals: importance of configurational thermodynamic quantities and molecular mobility. *J Pharm Sci* 2002;**91**(8):1863–72.
43. Zhou D, Schmitt EA, Zhang GG, Law D, Vyazovkin S, Wight CA, et al. Crystallization kinetics of amorphous nifedipine studied by model-fitting and model-free approaches. *J Pharm Sci* 2003;**92**(9):1779–92.
44. Zhou D, Grant DJW, Zhang GGZ, Law D, Schmitt EA. A calorimetric investigation of thermodynamic and molecular mobility contributions to the physical stability of two pharmaceutical glasses. *J Pharm Sci* 2007;**96**(1):71–83.
45. Buckton G. Applications of isothermal microcalorimetry in the pharmaceutical sciences. *Thermochimica Acta* 1995;**248**:117–29.
46. Jakobsen DF, Frokjaer S, Larsen C, Niemann H, Buur A. Application of isothermal microcalorimetry in preformulation. I. Hygroscopicity of drug substances. *Int J Pharm* 1997;**156**(1):67–77.
47. Pikal MJ, Dellerman KM. Stability testing of pharmaceuticals by high-sensitivity isothermal calorimetry at 25°C: cephalosporins in the solid and aqueous solution states. *Int J Pharm* 1989;**50**(3):233–52.
48. Schmitt EA, Peck K, Sun Y, Geoffroy J-M. Rapid, practical and predictive excipient compatibility screening using isothermal microcalorimetry. *Thermochimica Acta* 2001;**380**(2):175–84.
49. Findlay WP, Bugay DE. Utilization of Fourier transform-Raman spectroscopy for the study of pharmaceutical crystal forms1. *J Pharm Biomed Anal* 1998;**16**(6):921–30.
50. Seyer JJ, Luner PE, Kemper MS. Application of diffuse reflectance near-infrared spectroscopy for determination of crystallinity. *J Pharm Sci* 2000;**89**(10):1305–16.
51. Chan KLA, Fleming OS, Kazarian SG, Vassou D, Chryssikos GD, Gionis V. Polymorphism and devitrification of nifedipine under controlled humidity: a combined FT-Raman, IR and Raman microscopic investigation. *J Raman Spectros* 2004;**35**(5):353–9.
52. Basavoju S, Boström D, Velaga SP. Pharmaceutical cocrystal and salts of norfloxacin. *Crystal Growth Design* 2006;**6**(12):2699–708.
53. Srivastava HK, Wolfgang S, Rodriguez JD. Expanding the analytical toolbox for identity testing of pharmaceutical ingredients: spectroscopic screening of dextrose using portable Raman and near infrared spectrometers. *Anal Chim Acta* 2016;**914**:91–9.
54. Clarke FC, Jamieson MJ, Clark DA, Hammond SV, Jee RD, Moffat AC. Chemical image fusion. The synergy of FT-NIR and Raman mapping microscopy to enable a more complete visualization of pharmaceutical formulations. *Anal Chem* 2001; **73**(10):2213–20.
55. Kazarian SG, Martirosyan GG. Spectroscopy of polymer/drug formulations processed with supercritical fluids: in situ ATR–IR and Raman study of impregnation of ibuprofen into PVP. *Int J Pharm* 2002;**232**(1–2):81–90.
56. Tobyn M, Brown J, Dennis AB, Fakes M, Gao Q, Gamble J, et al. Amorphous drug–PVP dispersions: application of theoretical, thermal and spectroscopic analytical techniques to the study of a molecule with intermolecular bonds in both the crystalline and pure amorphous state. *J Pharm Sci* 2009;**98**(9):3456–68.
57. Andrews GP, AbuDiak OA, Jones DS. Physicochemical characterization of hot melt extruded bicalutamide–polyvinylpyrrolidone solid dispersions. *J Pharm Sci* 2010;**99**(3):1322–35.
58. Reddy LS, Bethune SJ, Kampf JW, Rodríguez-Hornedo N. Cocrystals and salts of gabapentin: pH dependent cocrystal stability and solubility. *Crystal Growth Design* 2009;**9**(1):378–85.
59. Gowen AA, O'Donnell CP, Cullen PJ, Bell SEJ. Recent applications of Chemical Imaging to pharmaceutical process monitoring and quality control. *Eur J Pharm Biopharm* 2008;**69**(1):10–22.
60. De Beer T, Burggraeve A, Fonteyne M, Saerens L, Remon JP, Vervaet C. Near infrared and Raman spectroscopy for the in-process monitoring of pharmaceutical production processes. *Int J Pharm* 2011;**417**(1–2):32–47.
61. Jørgensen A, Rantanen J, Karjalainen M, Khriachtchev L, Räsänen E, Yliruusi J. Hydrate formation during wet granulation studied by spectroscopic methods and multivariate analysis. *Pharm Res* 2002;**19**(9):1285–91.
62. Ricci C, Eliasson C, Macleod NA, Newton PN, Matousek P, Kazarian SG. Characterization of genuine and fake artesunate anti-malarial tablets using Fourier transform infrared imaging and spatially offset Raman spectroscopy through blister packs. *Anal Bioanal Chem* 2007;**389**(5):1525–32.
63. Deisingh AK. Pharmaceutical counterfeiting. *Analyst* 2005;**130** (3):271–9.
64. Sacré P-Y, Deconinck E, De Beer T, Courselle P, Vancauwenberghe R, Chiap P, et al. Comparison and combination of spectroscopic techniques for the detection of counterfeit medicines. *J Pharm Biomed Anal* 2010;**53**(3):445–53.
65. Hedoux A. Recent developments in the Raman and infrared investigations of amorphous pharmaceuticals and protein formulations: a review. *Adv Drug Deliv Rev* 2015. <http://www.sciencedirect.com/science/article/pii/S0169409X15002859>
66. Wartewig S, Neubert RHH. Pharmaceutical applications of Mid-IR and Raman spectroscopy. *Adv Drug Deliv Rev* 2005; **57**(8):1144–70.

67. Gendrin C, Roggo Y, Collet C. Pharmaceutical applications of vibrational chemical imaging and chemometrics: a review. *J Pharm Biomed Anal* 2008;**48**(3):533–53.
68. Duer MJ. *The Basics of Solid-State NMR*. Oxford: Blackwell Science Ltd; 2001.
69. Levitt MH. *Spin dynamics: basics of nuclear magnetic resonance*. Chichester: John Wiley & Sons, Ltd; 2002.
70. Gunther H. *NMR spectroscopy: basic principles, concepts and applications in chemistry*. 3rd ed. Weinheim: WILEY-VCH; 2013.
71. Conte P, Piccolo A. Solid state nuclear magnetic resonance spectroscopy as a tool to characterize natural organic matter and soil samples. The basic principles. *ÓPTICA PURA Y APLICADA* 2007;**40**(2):215–26.
72. Bugay DE. Characterization of the solid-state: spectroscopic techniques. *Adv Drug Deliv Rev* 2001;**48**(1):43–65.
73. Byrn SR, Xu W, Newman AW. Chemical reactivity in solid-state pharmaceuticals: formulation implications. *Adv Drug Deliv Rev* 2001;**48**(1):115–36.
74. Harris RK. NMR studies of organic polymorphs & solvates. *Analyst* 2006;**131**(3):351–73.
75. Stephenson GA, Stowell JG, Toma PH, Pfeiffer RR, Byrn SR. Solid-state investigations of erythromycin a dihydrate: structure, NMR spectroscopy, and hygroscopicity. *J Pharm Sci* 1997;**86**(11):1239–44.
76. Nunes TG, Viciosa MT, Correia NT, Danède F, Nunes RG, Diogo HP. A stable amorphous statin: solid-state NMR and dielectric studies on dynamic heterogeneity of simvastatin. *Mol Pharm* 2014;**11**(3):727–37.
77. Booy K-J, Wiegerinck P, Vader J, Kaspersen F, Lambregts D, Vromans H, et al. The use of 13C labeling to enhance the sensitivity of 13C solid-state CPMAS NMR to study polymorphism in low dose solid formulations. *J Pharm Sci* 2005;**94**(2):458–63.
78. Yuan X, Sperger D, Munson EJ. Investigating miscibility and molecular mobility of nifedipine-PVP amorphous solid dispersions using solid-state NMR spectroscopy. *Mol Pharm* 2014;**11**(1):329–37.
79. Thijs HML, Becer CR, Guerrero-Sanchez C, Fournier D, Hoogenboom R, Schubert US. Water uptake of hydrophilic polymers determined by a thermal gravimetric analyzer with a controlled humidity chamber. *J Mater Chem* 2007;**17**(46):4864–71.
80. Nunes C, Han J, Monkhouse DC, Suryanarayanan R. Preformulation studies to meet the challenges in the manufacture of betaine solid dosage form. *J Pharm Sci* 2004;**93**(1):38–47.
81. Chrzanowski F. Preformulation considerations for controlled release dosage forms: part I selecting candidates. *AAPS PharmSciTech* 2008;**9**(2):635–8.
82. Braun DE, Karamertzanis PG, Arlin J-B, Florence AJ, Kahlenberg V, Tocher DA, et al. Solid-state forms of β-resorcylic acid: how exhaustive should a polymorph screen be?. *Crystal Growth Design* 2011;**11**(1):210–20.
83. Braun DE, Koztecki LH, McMahon JA, Price SL, Reutzel-Edens SM. Navigating the waters of unconventional crystalline hydrates. *Mol Pharm* 2015;**12**(8):3069–88.
84. Hancock KE, Peck GE, Perry DL, White JL, Hem SL. The effect of humidity on the physical and chemical stability of spray-dried aluminum hydroxycarbonate. *J Colloid Interface Sci* 1996;**183**(2):431–40.
85. Mahlin D, Berggren J, Alderborn G, Engström S. Moisture-induced surface crystallization of spray-dried amorphous lactose particles studied by atomic force microscopy. *J Pharm Sci* 2004;**93**(1):29–37.
86. Khoo JY, Shah UV, Schaepertoens M, Williams DR, Heng JYY. Process-induced phase transformation of carbamazepine dihydrate to its polymorphic anhydrates. *Powder Technol* 2013;**236**:114–21.
87. Waterman KC, MacDonald BC. Package selection for moisture protection for solid, oral drug products. *J Pharm Sci* 2010;**99**(11):4437–52.

CHAPTER

4

API Solid-Form Screening and Selection

M.Y. Gokhale and R.V. Mantri

Drug Product Science & Technology, Bristol-Myers Squibb, New Brunswick, NJ, United States

4.1 INTRODUCTION

Most active pharmaceutical ingredients (APIs) are solid in nature and could exist in many forms that vary significantly in their properties leading to different physical, chemical, mechanical, and biopharmaceutical properties.[1–6] The ideal API solid form is one that, not only can be manufactured reproducibly, characterized in detail with respect to identity and stability, but also be formulated as a drug product (DP) that meets the intended target product profile. Therefore, screening and selection of the API solid form are critical components in drug development that influence the quality, performance, and manufacturing of a drug substance and DP. Additionally, in today's competitive pharmaceutical landscape, since development timelines are getting much shorter, the need to select the right candidate the first time is becoming more important.

Solid forms of APIs can exist as crystalline forms including polymorphs, hydrates, solvates, cocrystals, and salts as well as amorphous forms. The aim of solid form screening and selection is to identify and characterize as many forms as possible and select a form with optimal physical, chemical, and biopharmaceutical properties.[7] Crystalline forms are generally preferred over amorphous forms because they can be reproducibly manufactured, often with greater purity and stability. Therefore, the solid form screening is primarily focused on the identification and characterization of crystalline forms prior to selection. The amorphous forms are selected either because a crystalline form is not identified or the known crystalline forms have inadequate solubility, dissolution rate, and bioavailability.

Typically API solid form selection strategy is executed by implementing a tiered or integrated approach that involves multiple steps.[7–9] It is desirable to identify a wide range of crystalline polymorphs, hydrates, and the relevant solvates during the form screening stage. Initial assessments that influence the selection strategy include considerations such as ionization of the molecule, solubility, and stability along with formulation considerations and route of administration.[10–16] For ionizable drugs (anionic, cationic, and zwitterionic), salt formation is the most common strategy to enhance solubility and bioavailability.[16,17] Salts can also provide improved solid-state properties, stability, crystallinity, hygroscopicity, melting point and bulk physical properties. For nonionizable drugs, weak bases, or acid molecules where salt formation is limited, cocrystals are an attractive option.[15,18] Screening of solid forms may be achieved by high-throughput and/or manual crystallization approaches with the goal of identifying crystalline forms using various cosolvents to assess solvates, hydrates, and potential counter ion screens for salts or cocrystals followed by scaling up of the crystalline leads. The lead candidates are characterized using solid-state analytical techniques such as thermal analysis, crystallography, X-ray diffraction (XRD), and spectroscopy.[19–23] An additional goal of form characterization is also to determine the thermodynamic relationship between known polymorphs (monotropic vs enantiotropic), pseudopolymorphs (eg, hydrate/anhydrate), salts (eg, pHmax for salt vs free form), or cocrystals (eg, K_{sp} and Keu). Assessment of hygroscopicity, solubility, and chemical and physical stability is also routinely conducted. Additional factors that influence the form selection include considerations for API isolation (crystallization, filtration, and drying) process as well as the formulation strategy to meet the target product profile. Thus, the selection of the form is done after careful review of results from characterization studies including physical, chemical stability, and solubility/dissolution behavior as well as in vivo

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00004-2

performance and manufacturing considerations. It is known that mechanical and physical property differences between forms can affect manufacturability and physical attributes of the solid oral dosage forms as has been reported for theophylline,[6] sulfamerazine,[24] polymorphs of metoprolol tartrate,[25] paracetamol,[26] or carbamazepine.[27] The mechanical properties are also influenced by particle size and shape/morphology in addition to the solid form. Generally, these properties are often optimized by particle engineering approaches on the selected form at a later stage in development and are outside the scope of this chapter.

4.2 SOLID-FORM SELECTION CONSIDERATIONS

Ideally, the selected form is a single-phase, nonhygroscopic, highly crystalline form. It also possesses excellent thermal, physical, and chemical stability to ensure good quality during manufacturing and storage of API and dosage form. The selected form does not have complexities of polymorphism and pseudopolymorphism and can be manufactured in large scale with good yield and purity in a consistent fashion. Although mechanical properties of the solid form are not thoroughly evaluated during the solid-form selection stage, the selected form should be amenable to shape/morphology and size as required by the selected formulation and process. Finally, the solid form should have the solubility and dissolution rate characteristics to provide pharmacokinetics and bioavailability as dictated by the target product profile of the dosage form.

4.2.1 Key physicochemical property considerations

4.2.1.1 Solid-form stability

Stability testing of drug substance and DP are conducted during development with a goal of establishing and assuring that quality is maintained during the recommended storage conditions and time (shelf-life). It is critical to select a solid form that is stable as a function of time and environmental factors such as temperature, humidity, and light. For APIs that exist in multiple solid-state forms, it is important to understand the relative thermodynamic stability of the various forms and its relevance not only to manufacturing but also to storage conditions and shelf-life.

Physical stability of solid forms is related to phase transformations and interconversion between different crystalline polymorphic or pseudopolymorphic forms or conversion of thermodynamically unstable amorphous to crystalline form or salt form to free form. As part of the form selection process, understanding these transformations in relevance to thermodynamic or kinetic factors during API processing or DP processing and on shelf-life is an important element of the selection process. Solid forms can display a wide variety of physicochemical properties, including solubility, melting points, density, and morphology, and transformation of one solid form to another form can lead to formation of undesirable forms and pose risks to quality and in-vivo performance. For example, for amorphous API or dispersions, maintaining the glassy state and keeping the material amorphous are critical to ensure higher solubility and improved bioavailability as intended with these systems. On the other hand, for crystalline forms, even small amounts of amorphous can result in undesired effects such as increased water uptake and reduced chemical stability. Similarly for salts or cocrystals, dissociation to the free form during storage or in the presence of certain formulation excipients can lead to poor performance. Various processing conditions during DP manufacture, including wet granulation, milling, drying, and compaction, can cause phase transformations from anhydrate to hydrate, crystalline to amorphous, desolvation of a solvate, or conversion to different polymorphs, and can all lead to lower solubility, slower dissolution rates, and potentially lower in vivo exposures.[28–31] It is, therefore, important to first understand the thermodynamic stability of known polymorphs, hydrates, and relevant solvates. The relative stability of polymorphs (ie, enantiotropic vs monotropic systems) needs to be assessed to ensure that the chosen form is thermodynamically stable. For salts, it is important to determine the relative stability of salt versus free forms as a function of pH to assess the potential risk(s) of dissociation in solid dosage form. As with salts, similar dissociation of cocrystals to individual components can lead to physical instability. Phase diagrams for cocrystals are useful in demonstrating conditions where a cocrystal component is supersaturated, and selection of these conditions should avoid dissociation of the cocrystal.[32] In case of hydrates and anhydrates, at a constant temperature over a range of given relative humidity (RH) or water activity, one crystal form is stable, and at a critical water activity, both forms can exist simultaneously. However, based on the kinetics and conditions of conversion, transient metastable forms may exist during the hydration or dehydration process. In aqueous systems, conversion to the more stable hydrate form is a key concern from a bioavailability standpoint, since hydrates are known to have lower solubility typically in the order of two- to threefold.[33] Additionally, conversion to hydrate on storage during shelf life can also impact the chemical stability of the DP. The anhydrate-to-hydrate conversion is an important consideration in designing

solution dosage forms to ensure that the hydrate solubility is not exceeded, and in the case of suspensions, no conversion is seen during the shelf life that can lead to increased particle size and caking of the suspension leading to instability. The extent and time taken for conversion to the more stable form can be controlled by modifying the kinetics, and it depends on various factors such as the solubility of a drug, energy input during mechanical processes, temperature, and the presence of certain excipients.[33]

While selection of the most stable form is important to ensure adequate physical stability, when the desired physicochemical properties are not achieved, the use of metastable forms with higher solubility such as amorphous solid dispersions are widely used as a strategy for improved in-vivo performance.[34] In such cases where the metastable forms are chosen for development, understanding the kinetic stability of these systems is extremely critical. Molecular mobility is an important consideration in understanding the amorphous-to-crystalline phase transitions, and this has been discussed later.

Another important stability consideration is the chemical stability that involves the degradation of an API or a reaction with excipients to form new degradants.[13] During form selection, evaluating the chemical stability of a selected form is an integral part of the form screening process. In some cases, polymorphs (or psuedopolymorphs) of some drugs can exhibit different chemical stability, as in the case of carbamazepine,[35] indomethacin,[36] or furosemide,[37] to name a few. Generally speaking, the thermodynamically stable form is the most chemically stable. However, in case of indomethacin, the metastable α-form has a higher density than the more stable γ-form, but it is still more reactive. This is due to differences in hydrogen bonding and crystal packing that expose the reactive carboxylic acid group on the crystal surface for the α-form.[36]

Chemical instability can also be seen as an outcome of physical changes in the form. For example, in the case of either amorphous forms or crystalline API where amorphous regions are introduced during the various processing operations such as grinding, milling, compaction, or lyophilization, molecular mobility plays an important role in chemical reactivity in the solid state. Increased molecular mobility is shown to result in increased rates of chemical degradation in a wide variety of reactions, including hydrolysis, cyclization, and even oxidation.[38,39] In such cases, water acts as a plasticizer and is important when considering such drugs or excipients. Amorphous material can sorb large amounts of water, which will be more localized in the amorphous regions,[40,41] leading to varying chemical reactivity depending on the extent of such disordered regions in the DP. Similarly for drug excipient interactions, water can play a significant role by increased water content brought in by the excipient that is adsorbed into the drug or increases the amount of water in the sorbed layer at the drug excipient point of contact, thus increasing or facilitating the rate of chemical degradation with that excipient.

Typical solid-state reactions include hydrolysis, oxidation, cyclization, and racemization. Excipients used in DPs are typically related to chemical instability by virtue of the reactive impurities found in these systems and the choice of form influences the potential chemical reactivity.[42] Similarly polymorphs can also have varying stability as seen in the case of chloroquine diphosphate,[43] where the stability under light varies, or for the various polymorphs of prednisolone *t*-butyl acetate, where only one form was found to be susceptible to oxidation and had a tunnel parallel to the hexagonal axis of the crystal that allowed oxygen to penetrate, while the other forms had closely tight packing.[44]

4.2.1.2 Hygroscopicity

Hygroscopicity refers to the ability of a solid to take up and retain water at various RH and temperatures. Both the equilibrium moisture and the rate of uptake are important to understand to assess the hygroscopicity of various forms and potential impact. Various forms, morphology, crystallinity, and the state of hydration affect the water uptake under different experimental conditions. Low hygroscopicity is always preferred for the selected API form. Based on the extent of water uptake, pharmaceutical solids can be classified into nonhygroscopic, slightly hygroscopic, and hygroscopic solids.[45] Moisture adsorption isotherms, plots of equilibrium water content of a solid as a function of RH at a constant temperature, are commonly used to assess the hygroscopicity of pharmaceutical solids.

Hygroscopicity can be determined by storing solids at various controlled RH and temperature over long periods of time to equilibrate and measure weight gain or more typically in recent times by the automated sorption microbalance technique or dynamic vapor sorption instruments. In this, small quantities of drug (about 5 mg) are purged with humid nitrogen and equilibrated at controlled humidity and temperature, and the weight change is recorded. Highly crystalline solids typically have negligible surface adsorption of water, whereas high amounts of water uptake are seen typically for amorphous or deliquescent materials. Stepwise increase in water uptake is typically seen for stoichiometric hydrates, while nonstoichiometric or channel hydrates show a more gradual increase in water uptake as a function of RH.[46–48] Typically, percent weight change associated with the solid can be

used to determine the stoichiometry of the hydrate based on the molecular weight of the compound. Measurements of hygroscopicity using the dynamic sorption technique typically include an initial drying phase, which should be avoided for forms that may be solvates or hydrates. Differences in kinetics of adsorption and desorption are sometimes observed, leading to the two isotherms not coinciding and resulting in a hysteresis. The most common reason for hysteresis is the slower sorption kinetics, but a few other likely reasons are deliquescence at high humidity followed by no recrystallization during the desorption phase or amorphous material undergoing conformational changes due to water uptake or formation of a new form (hydrates) that show varying kinetic stabilities.[49,50] Typically adsorption-desorption studies are coupled with analysis of the solids such as powder X-ray diffraction (PXRD) to confirm the API form identity at various parts of the adsorption-desorption cycle. Sometimes crystalline material due to surface amorphous regions show larger amounts of water uptake and can cause issues with processing and handling of pharmaceutical API during manufacture and storage. Measurement of hygroscopicity of early lots of API showing differences in moisture uptake profiles can be an indication of amorphous regions, phase impurities, and crystal defects and provides information on early lot-to-lot variability. These studies can provide important insight into phase transformations and critical water activity ranges between hydrates and anhydrate forms as long as the isotherms are completely reversible; however, when hysteresis is observed, the kinetics are critical in correlating relevance on long-term moisture uptake.[51]

Increased hygroscopicity can lead to an increase in molecular mobility and solid-state reactivity. Deliquescence, conversion of forms (formation of hydrates), and potential chemical instability due to higher degradation are all possible due to increased water absorption by surfaces of solids.[13] Additionally, physical instability due to changes in moisture content leading to dissociation of salts, conversion of hydrates to anhydrates or vice versa during DP processing or storage can result in dissolution slow down.[52,53] Given the impact of changes in moisture content of the API form to alter physicochemical properties and altered product performance, low hygrosopicity is a desirable property during form screening. Understanding the critical water activity ranges for various hydrates/anhydrates is also important. Considerations of the potential impact of unit operations such as drying, milling on API form, and the potential generation of crystal defects, amorphous regions, and altered moisture uptake profiles are important considerations during form evaluation.

The hygroscopicity behavior can vary significantly across various forms, polymorphs, or hydrates of a given molecule. For example, in the case of carbamazepine, which is hygroscopic at high RH and converts to the dihydrate, formation of cocrystals with nicotinamide coformer have shown to prevent dihydrate formation up to 1 month at 40°C/94% RH.[54]

4.2.1.3 Solubility, dissolution rate, and bioavailability

Differences in crystal packing result in wide range of physicochemical properties, including melting point differences, differences in internal energy, enthalpy, free energy, chemical potential, and solubility. In case of polymorphs or pseudopolymorphs, the thermodynamic stability also dictates the solubility. Pudipeddi et al.[55] surveyed over 55 compounds and their polymorphs/pseudopolymorphs and found that the average solubility ratio of the individual polymorphs to the most stable (least soluble) one was 1.7.

Polymorphs with two-fold or larger difference in solubility may have improved in-vivo performance, as seen in the case of chloramphenicol palmitate, where a fourfold solubility difference showed a sixfold difference in C_{max} in humans, whereas mefenamic acid with a 1.3-fold solubility difference showed no significant improvement in vivo.[1] Another classic case of the impact of polymorphs on solubility and dissolution rate is for ritonavir. A new thermodynamically stable form II was discovered a couple of years after form I was launched in the market. The two forms had a greater than fivefold solubility difference, and hence form II supersaturated and was much less soluble in the semisolid formulation, forcing the manufacturer to recall the original formulation.[56,57] However, the impact on bioavailability between various polymorphs also depends on additional factors such as solubility difference, dose, permeability, and formulation factors. Biopharmaceutics Classification System (BCS) analysis or computational simulations using physiologically based pharmokinetic (PBPK) modeling with software programs such as GastroPlus can be used to assess the interplay between solubility, dose, and permeability to predict overall impact on bioavailability.

The lower aqueous solubility typically imparted by hydrates is an important consideration when evaluating neat and hydrate forms for a given drug and its potential risk to in vivo exposure. The kinetic advantage of the higher aqueous solubility of the anhydrate form is important in the consideration of various forms or designing dosage forms that may impart longer supersaturation states with respect to the lower solubility hydrate forms.

Typical solubility measurements include the equilibrium solubility, which is the solubility at equilibrium. These measurements are usually run over 24 hours or longer, and the results can be impacted by changes in temperature during the study or phase changes such as the formation of hydrates, solvates, and disproportionation of salt forms. Evaluating solution-mediated phase transformation during these dissolution or solubility studies is an important consideration. Even though phase conversion is expected to occur based on thermodynamics driving force, the kinetics will govern whether it will occur during the relevant conditions and time. Assessing the kinetic dissolution as a function of time provides information not only on the equilibrium solubility but also the metastable forms and their dissolution rates. Two commonly used methods for assessing form solubility is powder dissolution and intrinsic dissolution.[58] In cases of polymorphs, salts, or cocrystals, if the form is physically unstable under biorelevant aqueous condition, equilibrium solubility may not be relevant to correlate to oral absorption, and kinetic solubility or intrinsic dissolution rates (IDRs) over a few hours are more relevant.

The intrinsic rate is a dissolution rate measured by an intrinsic apparatus and is independent of surface area. The rotating-disk system is used most commonly used, and a good correlation of results is observed compared with the fixed-disk system[59]. The IDR can serve as a surrogate to the equilibrium solubility studies. These measurements provide information on any phase transformations during the solubility study and are able to capture transient supersaturation concentrations. The IDR study results can vary based on internal and external factors (properties of solid state, surface area, hydrodynamic condition, composition of dissolution media, vibration, rotation speed, presence of dissolved gases in dissolution media, and sample sites), and hence when comparing results, the methods should be standardized. Any change in phase transition during the experiment such as the conversion of amorphous atorvastatin into crystalline,[60] conversion of diclofenac salt to free acid form,[2] or conversion of polymorphs can be picked up easily by nonlinear dissolution curves for the amount dissolved versus time due precipitation of the more stable form. However, intrinsic rates can still be obtained from the linear portion of the curves. The ratio of IDR of two solid forms of the same drug (eg, polymorphs, hydrates) corresponds to the solubility ratio between the same forms. Therefore, the IDR experiments are advantageous over solubility experiments when studying the relative stability of solid forms, especially when there is phase transformation during solubility studies.

Drug dissolution is the first step in the absorption process, and both solubility and the rate of dissolution play an important role in the in vivo absorption. The drug dissolution rate can be described by[61]:

$$DR = \frac{dX_d}{dt} = \frac{A \times D}{\delta} \times \left[\frac{C_s - X_d}{V}\right]$$

where A is effective surface area, D is the diffusion coefficient, δ is the diffusion layer thickness, C_s is saturation solubility, X_d is amount of drug in solution, and V is volume of dissolution media. In-vivo gastrointestinal conditions such as bile salt concentration and the pH level of the stomach can both impact the solubility or dissolution rate of the drug and affect in-vivo absorption.

When evaluating in-vitro dissolution or solubility measurements, it is important to have media that are close in composition to in-vivo conditions to have predictive ability. Typical aqueous buffers used as dissolution media can simulate the pH appropriately but cannot simulate the ionic strength, bile salt composition, surface tension, or viscosity. In order to simulate the in-vivo conditions adequately, assessing dissolution and solubility in more physiologically relevant in-vitro methods is important for poorly water-soluble drugs. The common biorelevant dissolution media are simulated gastric fluid and simulated intestinal fluid under fasted (FaSSIF) or fed (FeSSIF) conditions. Composition of the media greatly influences the predictability, and many studies have been conducted over the years to establish the compositions. For example, for FaSSIF and FeSSIF media, sodium taurocholate is used as representative bile salt since it closely represents cholic acid, which is found in human bile.[62] Additionally, the buffer capacity of FaSSIF is much lower than FeSSIF to represent in-vivo conditions. In the case of fed-state conditions, to represent the post-meal conditions, taurocholate and lecithin are included at much higher concentrations. A detailed review by Klein describes the various biorelevant media, compositions, and examples on the application of biorelevant dissolution and correlation to in-vivo studies for various drugs including danazol, itraconazole, or phenytoin.[61] As part of solid-form screening, especially for BCS Class II molecules where solubility is rate limiting, solubility measurements should be conducted in biorelevant media with dose/solubility ratios calculated, and dissolution tests in these media can potentially provide a tool for more meaningful in-vitro to in-vivo correlation.

Typically forms are designed to provide higher solubility and dissolution, which translate to higher bioavailability in vivo. Given the wide range of solubility and dissolution rates between forms, the choice of form can dictate bioavailability and can also impact

PK parameters such as C_{max} and T_{max}.[63] In vitro data generated in biorelevant media for various forms can serve as a selection criteria to pick lead candidates in preclinical animal studies to assess the impact on bioavailability. Assessing in-vivo performance of various forms in preclinical animal models such as dogs, monkeys, or rats under fed, fasted, or famotidine-treated conditions is an important aspect during selection. Data from in-vitro biorelevant solubility studies and preclinical animal PK data can then be used into computational simulations using physiology-based pharmacokinetic modeling softwares (eg, GastroPlus, Simulations Plus, Inc.) that are now routinely used to assess the impact of rate-limiting factors affecting the overall absorption. This approach can be very effective during form selection to identify limits on potential physicochemical properties such as solubility to maximize bioavailability. Chiang et al. demonstrated the use of PBPK modeling during the salt selection process of phenytoin to determine the solubility requirement for the salts at the relevant dose, which was used to guide the salt selection process.[64]

4.2.2 Considerations for various forms

As detailed in the previous section, the ideal solid form possesses desirable physical, chemical, and biopharmaceutical properties to enable both API and DP. The solid-form selection is conducted with this end goal in mind and involves identifying and characterizing multiple forms ranging from the crystalline solids (polymorphs, cocrystals, solvates/hydrates) to amorphous solids. The following section provides guidance on theoretical and practical considerations for salt and cocrystal formation, polymorphs/ hydrates, solvates, and amorphous solids.

4.2.2.1 Salts

A large fraction of drug molecules are ionizable and amenable for salt formation. Some of the key reasons to select salts over their free forms include increased solubility, dissolution rate, and/or bioavailability,[65] enhancement in physical and chemical stability, reduced hygroscopicity and improvement in purity, and handling and processability of API.[3,15–17,65–67] Salts may also afford improved powder properties such a compactibility leading to advantages in the DP manufacturing processes. One clear disadvantage of salt over the free form is the increase in the weight of a drug substance for the same amount of active moiety due to the inclusion of the counterion. In some cases, salt formation may be an additional step in the API manufacturing process. It is also possible that, for some cases, the salts may have inferior properties compared to the free forms.[68] Some of the key considerations in salt selection are detailed below.

4.2.2.1.1 pH-solubility profile and salt solubility

The aqueous solubility of weakly acidic and basic drugs is a function of pH and is governed by the ionization or dissociation constant.[17] The intrinsic solubility is defined as the solubility of the unionized form of the molecule. The relationship of pH and solubility of an ionizable molecule is described by the Henderson–Hasselbalch equation. The pH–solubility profile of a weak base can be described by two independent curves, one where the base is the equilibrium species, and the other where the salt is the equilibrium species. Fig. 4.1 shows the pH solubility of a weak base where it can be divided into Region 1 and 2. At equilibrium, at all pH values other than at the pH_{max} (pH of maximum solubility), only one solid phase (free form or salt) exists whatever is the starting material.

The total solubility for a base is a sum of its unionized (free base) and ionized form:

$$S = \left[BH^+\right] + [B]$$

where S is total solubility, $[BH^+]$ is the ionized base (salt), and [B] represents the unionized base.

As shown in Fig. 4.1, for a monobasic compound when dissolved in water, the following equilibrium exists:

$$BH^+ + H_2O \underset{K_a}{\Leftrightarrow} B + H_3O^+$$

or

$$K_a = \frac{[B][H_3O^+]}{[BH^+]}$$

where K_a is the ionization constant.

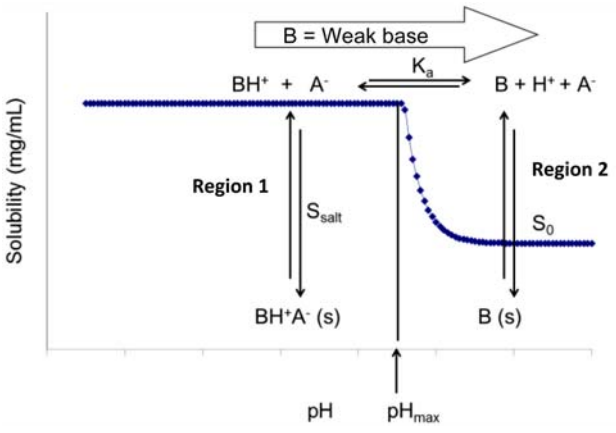


FIGURE 4.1 pH–solubility curve for a weak base.

In Region 2, where pH > $pH_{max,}$ the excess solid is the free base, and the solubility curve is described by the free base solubility and can be defined as

$$S = S_0\left(1 + \frac{[H_3O^+]}{\lfloor K_a \rfloor}\right)$$

where S is total solubility, S_0 is intrinsic solubility, and K_a is the ionization constant.

In Region 1 of Fig. 4.1 for a weak base, where pH < pH_{max}, the excess solid in equilibrium with the saturated solution is the salt. The solubility in this region is defined as:

$$S = \sqrt{K_{sp}}\left(1 + \frac{K_a}{[H_3O^+]}\right)$$

$$K_{sp} = [BH^+][A^-]$$

where K_{sp} is solubility product of the salt.

The pH value where Region 1 and 2 curves intersect is called the pH_{max} or the pH of maximum solubility.[69]

For a weak base, as the pH increases up to the $pH_{max,}$ the solubility is that of the salt. Even though the Figure 4.1 shows a pH independent region, there is a weak dependence on pH due to ionization of the salt. At pH values beyond the pH_{max}, the solubility decreases drastically due to decreased dissociation of salt to free base. At pH ≫ pK_a, the solubility again represents the intrinsic free base solubility and is independent of pH.

At the pH_{max}, in the presence of excess solid, the solution is saturated with respect to both the salt and the free form. Equating equations 2 and 3 above, the pH_{max} for bases can be defined as:

$$pH_{max} = pK_a + \log\frac{S_0}{\sqrt{K_{sp}}}$$

Similarly, the pH_{max} for weak acids is defined as:

$$pH_{max} = pK_a - \log\frac{S_0}{\sqrt{K_{sp}}}$$

The location of pH_{max} provides an indication of relative stability of salt versus free base as a function of pH in aqueous solutions or in the diffusion layer of solid particles, which provides the microenvironmental pH. The equation above shows the impact of the pK_a, $S_{0,}$ and K_{sp} on the pH_{max} and thus provides a measure of the stability of a salt as a function of aqueous pH.[70] For example, for a weak base, as the strength of the weak base increases (pK_a increases by one unit), the pH_{max} increases by one unit. Similarly, an order of magnitude increase in solubility of salt ($\sqrt{K_{sp}}$) decreases the pH_{max} by one unit. Given this, it can be perceived that the pH_{max} can shift over several orders of magnitude depending on the choice of counterions, which in turn relates to differences in the solubility or K_{sp} for the various salts and can be correlated to the melting point (crystal lattice forces).[71] In the case of weak acids above the pH_{max} and in the case of weak bases below the $pH_{max,}$ the solubility profiles are those of the respective acid or base salts. Ideally, the pH-solubility profiles generated using the free form or its salt should be superimposable when the same acid or base is used for pH adjustment in both cases. However, it is not uncommon to observe deviations from the ideality, as reported by Ledwidge and Corrigan.[72]

4.2.2.1.2 Selection of counterions and salt formation

Commonly used anions and cations typically considered for salt formation are listed in Table 4.1. Based on experience, for weakly basic acids and bases, a difference of two pK_a units (ie, the acid pK_a should be at least two units lower than the base) is ideal for obtaining a stable salt. This rule of thumb has not been formally proven but is often used in the salt selection.[14,73–75] Cruz-Cabeza[76] reviewed the relative occurrences of 6465 neutral and ionized acid-base crystal structures in the Cambridge Structural Database. Salts or ionized acid-base complexes were exclusively observed when the pK_a difference was 4 or greater, and nonionized acid-base complexes (cocrystals) were exclusively observed when the pK_a differences were −1 or lower. Both salts and cocrystals occurred when the pK_a difference was in the range of −1 and 4, although the probability of salt formation increased with an increase in pK_a difference in a linear fashion. Given this rule of thumb, very weak bases (pK_a 1−2) are limited to strong acids, whereas bases such as imidazole (pK_a 6−7) can form salts with a wider range of acids.[16] Paulekuhn et al.[77] surveyed the Orange Book database published by the Food and Drug Administration (FDA) in 2007 for the frequency of occurrence of different counterions in pharmaceutical salts. Chloride and sodium counterions were the most frequent for anionic and cationic choices respectively, but there was an increasing number of other counterions also being used.

A major factor in salt selection is enhancement of solubility and dissolution of oral solids, especially for immediate-release products. A priori prediction of salt solubility (K_{sp}) based on salt former properties such as hydrophilicity, pK_a is difficult. For example, as shown by Hsieh et al. in Table 4.2 for miconazole salts, despite of having a similar pK_a value to methanesulfonic acid (pK_a −1.89), the nitrate salt of miconazole (pK_a of nitric acid = −1.3) exhibits almost a threefold lower solubility.[78]

The salt solubility will not only depend on the structure and properties of the drug and salt former but also on the solid-state properties of the resulting

TABLE 4.1 Commonly Used Acids and Bases as Salt Formers

Counterion	pK_a,1	pK_a,2	pK_a,3
ACIDS			
Hydroiodic	−8		
Hydrochloric	−6		
Hydrobromic	−6		
Sulfuric	−3	1.92	
p-Toluenesulfonic	−1.34		
Nitric	−1.32		
Methanesulfonic	−1.2		
Maleic	1.92	6.23	
Phosphoric	1.96	7.12	12.32
L(+)-Tartaric	3.02	4.36	
Fumaric	3.03	4.38	
Citric	3.13	4.76	6.40
D,L-Lactic	3.86		
Succinic	4.19	5.48	
Acetic	4.76		
BASES			
Sodium	14		
Potassium	14		
Zinc	About 14		
Choline	13.9		
L-Arginine	13.2	9.09	2.18[a]
Calcium	12.7		
Magnesium	11.4		
Diethylamine	10.93		
L-Lysine	10.79	9.18	2.16[a]
Ammonium	9.27		
Meglumine	8.03		
Tromethamine	8.02		

[a]pK_a *of the carboxylic function of the amino acids.*
Adapted from Stahl PH, Sutter B. Salt Selection. In: Hilfiker R, editor. Polymorphism: in the Pharmaceutical Industry. Published Online: 22 May 2006 DOI: 10.1002/3527607889.ch12.

salt. The salt-forming counterions could impact salt solubility by changing both the crystal lattice and solvation energies[79,80]; therefore, salt solubility is often determined experimentally. Sometimes the experimental determination of aqueous solubility of salts may be challenging due to disproportionation (eg, for low pHmax), self-association, supersaturation, or other experimental factors.[70,79] An additional factor that can contribute to changes in salt solubility is the presence of common ions in solution. For example, the presence of sodium chloride led to a decrease in hydrochloride salt solubility compared to the theoretical curve due to the common ion (chloride) effect.[81]

Tong and Whitesell[75] developed an in situ salt-formation methodology that also provides a rapid assessment of solubility values of possible salts. A negative result of not observing a crystalline salt from an in situ salt-screening approach may not necessarily mean that a salt can't be crystallized. A more detailed approach to solid-form screening is discussed in a later section of this chapter. In the event that a crystalline salt is observed during the in situ screening approach, it allows for an estimation of salt solubility by combining mass balance and ionic equilibria equations. For determining solubility of the salt of a freebase, the approach involves measurement of solubility of weak bases in the presence of excess amounts of acids that generate salt-forming counterion and tracking the amount of base, the concentration of the acid, volume, and final pH. The solubility product for a 1:1 salt precipitated during the experiment is given by the following equation:

$$K_{sp} = \frac{S}{MW}\left[\frac{[A] - \frac{(X - S \times V)}{V \times MW}}{1 + [H^+]/K_a}\right]$$

where X (mg) is the amount of solid base added, S (mg/mL) is the solubility in the acid solution, MW is the molecular weight of the base, V (mL) is the volume of the solution, [A] is the molar concentration of the acid, pK_a' is the ionization constant of the acid used, and pH is the final pH at the end of the experiment. This is especially important for hydrochloride salts in vivo where the chloride ion common effect can result in further suppressing the hydrochloride salt solubility.

4.2.2.1.3 Dissolution and oral absorption of salts

Dissolution of ionizable drugs and salts in unbuffered and buffered systems has been explained by applying the diffusion boundary layer theory, taking into account the ionic equilibria.[82] The pH of the diffusion layer (sometimes referred to as the microenvironmental pH) is a key factor in controlling the dissolution rate of the solid. The diffusion layer pH for the free form of an ionizable drug could be very different from the bulk pH due to the self-buffering action of the free form and dependent on the pK_a and solubility of the free form.[82] Salts also have the self-buffering capability and thus influence the diffusion layer pH. A number of reports on the relative dissolution rates of salts and free forms in reactive media have been reported.[17,83]

TABLE 4.2 Physicochemical Properties of Miconazole Free Base and Salts

Salt	pK_a	pH_{max}	pH at saturation	Salt solubility (M)	Free base solubility (M)
Miconazole	6.9	–	–	–	2.4×10^{-6}
Miconazole camsylate	–	4.17	3.43	1.27×10^{-3}	–
Miconazole mesylate	–	1.44	1.47	6.88×10^{-1}	–
Miconazole nitrate	–	4.75	4.37	3.36×10^{-4}	–
Miconazole phosphate monohydrate	–	3.35	2.23	8.53×10^{-3}	–
Miconazole tosylate	–	4.44	3.63	6.97×10^{-4}	–

Adapted from Hsieh YL et al. Salt stability—the effect of pH_{max} on salt to free base conversion. Pharm Res 2015;32(9):3110–8.

TABLE 4.3 Intrinsic Dissolution Rates of Haloperidol Free Base, Haloperidol Hydrochloride, and Haloperidol Mesylate as a Function of pH

	Freebase		Chloride salt		Mesylate salt	
pH of dissolution medium (adjusted using HCl or NaOH)	Dissolution rate, mg/(min · cm^2)	Diffusion layer pH	Dissolution rate, mg/(min · cm^2)	Diffusion layer pH	Dissolution rate, mg/(min · cm^2)	Diffusion layer pH
1.1	0.032	1.11	0.025	0.88	0.033	1.04
2.0	0.246	4.76	0.155	1.85	0.865	1.91
3.1	0.061	5.93	0.292	3.01	2.037	2.61
5.0	0.002	7.00	0.291	4.89	1.962	3.02
7.0	–	–	0.157	4.85	1.760	2.99

Reprinted with permission from Shoufeng et al. Investigation of Solubility and Dissolution of a Free Base and Two Different Salt Forms as a Function of pH", Pharm. Res; 2005, Volume 22, Issue 4, pp 628–635.

Li et al.[84] reported that the dissolution rate dependence on dissolution medium pH for free base, chloride, and mesylate salts could be explained by the diffusion layer pH. A summary of the results are provided in Table 4.3. It is clear that at very low pH conditions, the free base and mesylate salt showed similar dissolution rates as they all converted to chloride salt. As the pH increased, the salts outperformed the free base; the mesylate salt showed significantly higher solubility relative to the chloride salt owing to its higher K_{sp} that is expected to lower the diffusion layer pH.

In this example, the diffusion layer pH for both the salts was lower than the $pH_{max,}$ and there was no conversion of the salt to free base. For many salts, that may not be the case when the diffusion layer pH favors the free form. Understanding the kinetics of supersaturation, dissolution, and precipitation during the dissolution process becomes very important. In solution, salts can form supersaturated solutions, which is a metastable kinetically favored process near the pH_{max} where solubility values greater than saturation or theoretical solubility of the salt are observed.[65,67,68,85,86] The ability to kinetically maintain supersaturation is an important factor in providing the dissolution rate advantage for salts. Hawley et al. outlined[87] two mechanisms of precipitation: bulk precipitation, where the free form precipitates in solution; and surface precipitation, where the free form precipitates on the surface of the dissolving salt. In bulk precipitation, the solution concentration exceeds a certain critical concentration only after which precipitation starts occurring and continues until the solubility reaches that of the free form. In surface precipitation, the free form starts precipitating on the surface of the salt, and since the free form has lower solubility, the dissolution curve immediately slows down. The surface precipitation mechanism impacts dissolution of solubility the most by forming a coating of the free form on the surface, decreasing its dissolution. For salts, the supersaturation is the highest at the surface or in the diffusion layer surrounding the particles, and this leads to faster rates of free form precipitation at the surface, causing a layer of free base and slowing down overall dissolution. Diffusion layer-based dissolution models have been described in literature to predict the surface pH in dissolution media.[82] Understanding these dissolution and precipitation mechanisms can enable formulation design by altering the diffusion layer by inclusion of excipients to modify the pH of the diffusion layer and thus increase dissolution and absorption.[88]

It is important to consider the pH changes as the drug transits through the gastrointestinal tract. Following dissolution of the salt, precipitation and redissolution may occur, depending on the properties of the drug including pH-solubility profile, absorption rate, dose, and particle size. These factors play a role in the choice of salts versus their free forms. Ku[89] proposed that for highly soluble drugs (ie, BCS Class I and III), salts may not be beneficial compared to the free forms from a purely oral absorption perspective. The importance of salts in overcoming dissolution and absorption problems for poorly soluble drugs can't be overstated. When salts of free bases are administered, dissolution readily takes place in acidic conditions in the stomach. The chloride common-ion effect may sometimes be observed with hydrochloride salts, or the drug may convert to hydrochloride salt in the solid phase. Upon transiting into the intestine, the portion of the drug that is dissolved may remain in solution prior to absorption, or it may precipitate as free base. The undissolved portion of the salt is likely to convert to the free base or dissolve. The salt forms are generally beneficial for oral absorption of weak bases. Salts of weak acids are expected to undergo rapid dissolution and potentially precipitate as free acid in the stomach. The dissolved drug is absorbed in the intestine, and the undissolved drug undergoes dissolution prior to absorption. Although the case for salts of weak acids may not appear to be strong from an oral dissolution-absorption perspective, the kinetics of supersaturation, dissolution, and precipitation needs must be considered for the individual salt and free acid forms.

4.2.2.1.4 Toxicity of counterions

When a drug is administered in a solid dosage form, depending on the salt form dosed, the in-vivo solubility can be varied, thus influencing the pharmacokinetic and toxicological profile.[90] The choice of counterion can impact absorption through modification of the dissolution rate. The potential toxicity of a counterion is related to key factors such as dose, route of administration, pharmacokinetics, and the therapeutic indication.

Some counterions may lead to the formation of toxic impurities that may need to be controlled. For example, sulfonic acid salts such as mesylate, tosylates, and even sulfate salts can potentially lead to the formation of alkyl esters, which are known to be genotoxic in the presence of alcoholic solvents during processing. The toxicity concern can be addressed by implementing a good control strategy to monitor their levels in API and DPs.[91]

Pfannkuch et al.[10] has classified salt formers into three categories based on their toxicity profiles. Class 1 includes counterions such as chloride and sodium that are physiologically ubiquitous. Class 2 includes counterions that have been widely used and have relatively low toxicity and good tolerability, and Class 3 includes those that may be used for special considerations, for example, to provide some pharmacological effects.

In the case of cations, metal ions specifically such as sodium, potassium, calcium, and magnesium, since naturally abundant under physiological conditions, do not pose toxicity concerns and are considered inert at typical physiological levels. Similarly, nitrogen bases that are essential amino acids are of no concern. However, other cations such as lithium are used as therapeutic materials and are not inert. Also, long-term dosing of lithium in humans can cause irreversible damage to the kidneys,[92] while maleate salts have been reported to cause renal damage in dogs even after a single dose at 9 mg/kg.[93] Anions such as chlorides and phosphates are physiologically abundant and well tolerated, while other anions such as bromide, iodide, and nitrate need to be considered based on the amounts used. Certain salts may cause local irritation such as alprenolol HCl, which has an irritant effect on the esophagus, but benzoate salt has no such effect.[94] A more detailed review of toxicology considerations of various counterions and the effect of dose and route of administration for nonclinical and clinical use have been reported by Thackaberry[95] and Saal and Becker.[90] Tables 4.4 and 4.5, adapted from Saal and Becker, categorize the commonly available and maximum doses for various counterions for oral use. However, other considerations such as special populations (pediatric/geriatric/nursing women) should be considered if relevant during salt selection. Additional details on toxicology studies conducted for each pharmaceutical salt can be found on the FDA websites.

While the choice of a counterion during salt selection is mainly driven to enhance bioavailability or stability, safety of these salts is a key criteria of selection. Given this, while counterions are typically considered inert, along with factors such as ionization constants, solubility, and product stability, toxicity considerations are a key element in counterion selection. Other factors such as thermodynamic stability, disproportionation risk to free form, and supersaturation phenomenon shown by salts all go into the selection process and are discussed later.

4.2.2.1.5 Chemical stability considerations

The form of the API plays a significant role in influencing the chemical stability of a drug. Salt formation can lead to enhancement of the chemical stability of drugs. The microenvironmental pH also plays a significant role in determining the chemical reactivity of the DP. For example, for molecules that are acid labile, the selection of an acidic salt will increase the propensity

TABLE 4.4 Identified Maximum Daily Dose of Anionic Salt Formers Used in Oral Products

Salt former	Maximum daily dose (counterion)	Salt
Acetate	50 mg	Flecainide acetate
Benzoate	14 mg	Rizatriptan benzoate
Besylate	160 mg	Mesoridazine besylate
Bromide	300 mg	Pyridostigmine bromide
Carbonate	1500 mg	Lithium carbonate
Citrate	5250 mg	Piperazine citrate
Edisylate	20 mg	Prochlorperazine edisylate
Estolate	1450 mg	Erythromycin estolate
Fumarate	120 mg	Quetiapine fumarate
Gluconate	730 mg	Quinidine gluconate
Hippurate	1120 mg	Methenamine hippurate
Iodide	99 mg	Potassium iodide
Malate	60 mg	Diltiazem malate
Maleate	250 mg	Acetophenazine hydrogenmaleate
Mesylate	420 mg	Nelfinavir mesylate
Methylsulfate	200 mg	Diphemanil methylsulfate
Napsylate	170 mg	Propoxyphene napsylate
Oxalate	5 mg	Escitalopram oxalate
Pamoate	325 mg	Hydroxyzine pamoate
Phosphate	380 mg	Chloroquine dihydrogenphosphate
Stearate	1500 mg	Erythromycin stearate
Succinate	90 mg	Loxapine hemisuccinate
Sulfate	380 mg	Indinavir sulfate
Tartrate	3900 mg	Cysteamine hydrogentartrate
Tosylate	890 mg	Lapatinib ditosylate

Chloride is not included in this table since it is an endogenous physiologically ubiquitous ion.
Adapted with permission from Saal,C.; Becker,A. Pharmaceutical salts: A summary on doses of salt formers from the Orange Book, European Journal of Pharmaceutical Sciences 49 (2013) 614–623.

TABLE 4.5 Identified Maximum Daily Dose of Cationic Salt Formers Used in Oral Products

Salt former	Maximum daily dose (counterion)	Salt
Calcium	672 mg	Edetate calcium
Lithium	338 mg	Lithium carbonate
Lysine	860 mg	Ibuprofen lysine
Magnesium	4.9 mg	Omeprazole hemi-magnesium
Meglumine	8220 mg	Diatrizoate meglumine
Olamine	21 mg	Eltrombopag diolamine
Potassium	3575 mg	Aminosalicylate potassium
Tromethamine	2650 mg	Fosfomycin tromethamine

Sodium ion not listed since it's an endogenous and physiologically ubiquitous ion.
Adapted with permission from Saal, C.; Becker, A. Pharmaceutical salts: A summary on doses of salt formers from the Orange Book, European Journal of Pharmaceutical Sciences 49 (2013) 614–623.

of a hydrolysis reaction. Among the strategies to influence the chemical reactivity of a drug is through the modification of the microenvironmental pH. Badawy et al. showed that by choice of the mesylate salt form of the API or addition of acidic pH modifiers, which lower the microenvironmental pH, the ester hydrolysis reaction was successfully mitigated to enable a stable DP.[3,96] Similar to a stabilization effect, the use of excipients that can destabilize a salt form, as in the case of sertraline salts, was shown to induce conversion to free base in the DP leading to oxidative instability.[66] It is also possible that degradation may be exacerbated by salt formation either due to unfavorable microenvironmental pH conditions or direct interaction between the drug and counterion (eg, Michael addition between primary amine and fumaric acid).

4.2.2.1.6 Disproportionation of salts

In the case of salts, the physical stability related to the risk of disproportionation both in the solid state in DP and in vivo are key considerations, and kinetics of supersaturation and conversion to free form need to be assessed during salt selection. Disproportionation of salts is dissociation of the ionic species into its neutral and acidic or basic counterion, and this typically leads to the loss of the improved solubility due to formation of lower solubility free acid or base or alternately poor physicochemical properties, including lower stability. In case of cocrystals, dissociation involves the separation of the API molecules and the coformer.

The general rule of thumb to form stable salts is for the counterion and the active molecule pK_as to differ by a unit of 2. However, this is not the only criteria governing the disproportionation potential. Some critical factors include the microenvironmental pH (in solid state) and the pH_{max}. The microenvironmental pH is the pH of the microscopic layer that surrounds the solid particles of API, where a saturated solution of the solid exists due to adsorbed moisture. In oral solid dosage forms, the pH of the excipients that are mixed with the API dictate the microenvironmental pH.

Hence, modification of the microenvironmental pH of salts of drugs in solid dosage forms has been a successful strategy to mitigate the risk of disproportionation.[96,97] Similarly for acidic salts, the addition of basic excipients can increase the microenvironmental pH, thus resulting in conversion to free base, and depending on the reactivity of the various forms, the free base could be more reactive, leading to increased chemical degradation. The choice of excipients used in the DP will influence the microenvironmental pH, which then dictates the extent of disproportionation depending on the pH when compared against the pH_{max}.[66,98] If the microenvironmental pH is greater than the pH max, then the difference in solubility in the saturated layer between the free form and salt will govern the stability.

As described above, the pH solubility profile of a weak acid or base is a function of the K_{sp}, pK_a, and intrinsic solubility. For a weak base molecule, in the region above the $pH_{max,}$ the free base solubility is limiting, whereas in the region below the $pH_{max,}$ the salt solubility is limiting. At the $pH_{max,}$ the salt and acid are in equilibrium, and it represents the pH of maximum solubility, as described in the earlier section.

The microenvironmental pH will dictate whether thermodynamically disproportionation will be favored. Based on the equation for the pH_{max}:

$$pH_{max} = pK_a + \log \frac{S_0}{\sqrt{K_{sp}}}$$

where S_0 is solubility of the base, and $\sqrt{K_{sp}}$ is the intrinsic solubility of salt. Consistent with this equation, it can be seen that the more soluble a salt, the lower the $pH_{max,}$ resulting in a narrower practical operating pH range where the salt will be stable based on the microenvironmental pH. Thus, it's not only the pK_a that governs the disproportionation potential but also the pH_{max}. This is demonstrated in Fig. 4.2 by Hsieh et al., where a strong correlation is seen between the extent of disproportionation and the pH_{max} for a series of miconazole and sertraline salts.[78] Additionally, the buffering capacity of the salt may also play a role in dictating the disproportionation potential.[78,99]

Stephenson et al.[100] conducted a survey of around 203 active molecules and their counterions, which shows that no salts were reported for weak bases with pK_as of less than 4.6.

Along with these concepts, another important factor is the buffering capacity of the salt. As expected, salts with a lower buffering capacity tend to disproportionate more. The choice of excipients, for example magnesium stearate or croscarmellose sodium, and their buffering capacity influences the microenvironmental pH and are also important in the disporportionation potential.[99,101]

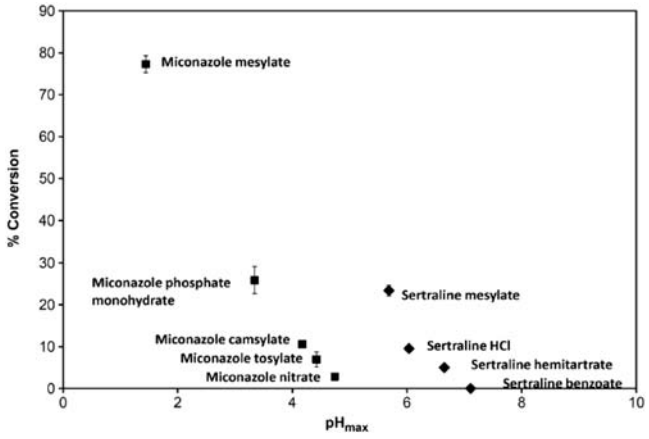


FIGURE 4.2 Relationship between pH_{max} and the extent of disproportionation for a series of miconazole and sertraline salts. *Reproduced from Hsieh YL et al. Salt stability—the effect of pH_{max} on salt to free base conversion. Pharm Res 2015;32(9):3110–8 with permission.*

This dependency on pH_{max} has led to building a theoretical model to predict the salt disproportionation[78,102] in the solid state and correlates to experimental data showing reasonable success. The key assumption for this model was a solution-mediated conversion, and the accuracy of predictions was found to be related to water content for the various salts and the buffering capacity provided by the API itself. This figure shows how the moisture content influences the equilibria for salt disproportionation:

$$[BH^+A^-]_s \xrightleftharpoons{\text{Sorbed } H_2O} [BH^+]_{aq}$$

$$+ [A^-]_{aq} \xrightleftharpoons[\text{Salt formation } pH < pH_{max}]{\text{Disproportionation } pH > pH_{max}} [B]_{aq}$$

$$+ [H^+]_{aq} + [A^-]_{aq} \rightleftharpoons [B]_s$$

This relationship to pH_{max} also highlights that while having a salt with high solubility is typically desirable, the salt with highest solubility may not be the best form,[103] but rather the optimal salt balancing the solubility with the disproportionation risk should go into selection of a salt form.

Additional factors such as temperature, particle size, RH, and drug-to-excipient ratios can also influence the extent of disproportionation. Studies by Hsieh at al. demonstrate that for three mesylate salts studied, disproportionation kinetics in the presence of a base at high humidity (57%) showed the disproportionation curve to overlap with the weight gain curve, which was indicative of the moisture uptake by the salts during the study course.[78,101] Similarly, an increase in surface area and a lower drug-to-excipient ratio also results in higher disproportionation, which is likely related to

an increased surface area per unit mass that results in increased water uptake or higher excipient contact, resulting in conversion to free form. The effect of temperature is opposite to that of moisture uptake, that is, an increase in temperature causes reduced moisture uptake but results in increased disproportionation.[101] The increased propensity of disproportionation with temperature is related to an increase in solubility of a given salt, which results in a subsequent decrease in the pH_{max}. This discussion highlights the importance of formulation composition as well as process and storage conditions considerations in the successful development of salts of weak acids and bases.

4.2.2.1.7 Dosage form consideration

The selection of any API solid form, including salt selection, needs to be done with the DP in mind. The considerations of dose and dosage form are critical in salt-form selection. Dose is an important consideration in the development of drug candidates and is exemplified in the BCS. It is possible that for some poorly soluble ionizable drugs salt form is not advantageous over the free form at low doses but could become beneficial at higher doses. Such an example is presented as Case Study 1 in this chapter. The solid form requirements for a controlled-release formulation is different from the immediate-release dosage form. Low solubility may be advantageous for controlled or extended-release formulations, and, therefore, poorly soluble salts may be advantageous. As described in the above sections, the choice of salt form of weakly ionizable drugs will likely influence the selection of excipients and formulation composition as well as the DP process.

4.2.2.2 *Cocrystals*

In recent years, similar to salts, another class of molecules called cocrystals has advanced significantly due to its ability to enhance solubility and dissolution or to crystallize otherwise difficult-to-crystallize drugs (eg, amorphous drugs with pK_as not favorable for salt formation). Fig. 4.3 depicts potential configurations and molecular differences between cocrystals and salts. A cocrystal is a molecular multicomponent complex that contains two or more different molecules in the same crystal lattice.[104] Cocrystals differ from solvates in that all components are solids at room temperature and from salts based on the extent of proton transfer. Cocrystals depend on noncovalent, nonionic interactions, which include hydrogen bonding, $\pi-\pi$, and van der Waals interactions. The properties of cocrystals can be varied significantly by the choice of coformer and component stoichiometry. Similar to salts, the physicochemical properties for cocrystals including

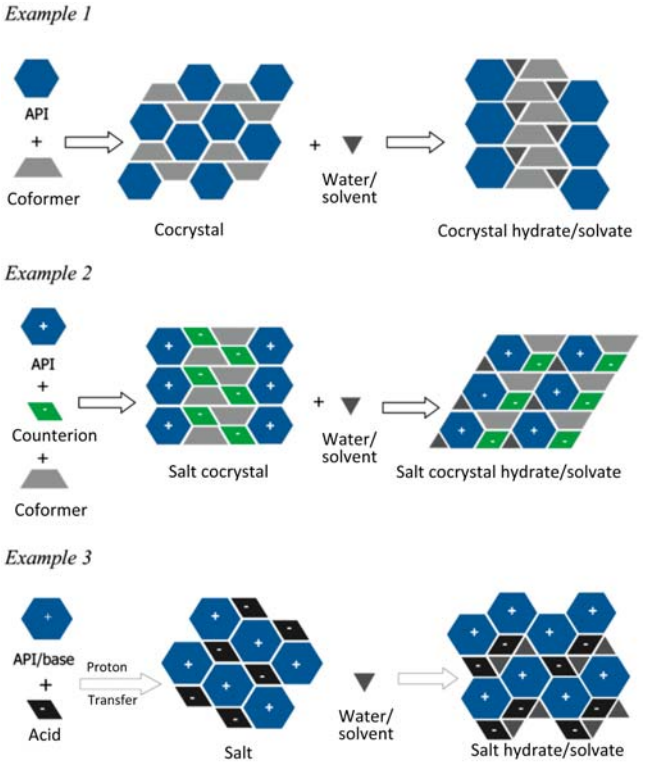


FIGURE 4.3 Possible multicomponent systems: cocrystals, salt cocrystals, and salts along with their respective solvate/hydrate forms. *Reproduced from Schulthesiss N, Newman A. Pharmaceutical cocrystals and their physicochemical properties. Crystal Growth Design 2009;9(6):2950–67 with permission.*

solubility, stability, dissolution rate, and melting point are important to assess during form screen.

Enhancement of solubility is the one of the primary drivers for cocrystal formation. As with salts, solubility is defined based on the solubility product, K_{sp}. For a cocrystal where A is a drug and B is a ligand, the solubility is described by the chemical equilibrium of solid cocrystal with solution and the corresponding equilibrium constant given by[105,106]:

$$A_\alpha B_{\beta(s)} \rightleftharpoons \alpha A_{(soln)} + \beta B_{(soln)}$$

Subscripts refer to the stoichiometric number of molecules of A or B in the complex. The equilibrium constant is given by

$$K \rightleftharpoons \frac{a_A^\alpha a_B^\beta}{a_{AB}}$$

If the activity of the solids is equal to 1, and assuming the activity coefficients (γ) of A and B equal unity for low-solute levels, cocrystal solubility can be described by solubility product K_{sp}:

$$K_{sp} = a_A^\alpha a_B^\beta \approx [A]^\alpha [B]^\beta$$

where [A] and [B] are the concentrations of components.

Two key factors influence the solubility behavior of cocrystals. One is the crystal lattice energy and solvation of the cocrystal components. Decreasing the crystal lattice energy or increasing the solvent affinity both lead to increased solubility of the cocrystal. In the case of cocrystals, a poor correlation is seen of solubility to melting point, but a reasonable correlation is seen with coformer solubility, indicating that solvation is often a greater parameter compared to lattice energy.[105,107]

4.2.2.2.1 Selection of coformer

The science of crystal engineering to form cocrystals is growing, and a number of API characteristics including hydrogen donor acceptor ability, pK_a, conformational flexibility, and solubility are important factors. An extensive review of cocrystals and coformers that existed prior to 2000 is compiled in the article by Stahly.[108]

The solubility values of 25 cocrystals from combinations of three drugs (carbamazepine, theophylline, and caffeine) and seven coformers (malonic acid, nicotinamide, salicylic acid, saccharin, succinic acid, glutaric acid, and oxalic acid) in four solvents (water, isopropyl alcohol, methanol, and ethyl acetate) were studied and ranked based on their solubility advantage over the drug.[26] The measured cocrystal solubility value ranged from 0.1 to over 100-fold of their respective drug solubility, and the coformer solubility spanned several orders of magnitude. A plot of cocrystal-to-drug solubility ratio against coformer-to-drug solubility ratio for each cocrystal showed a linear dependence. The larger coformer-to-drug solubility ratios showed a higher cocrystal solubility relative to the drug. The work demonstrated that cocrystal solubility enhancement could be rationally selected based on knowledge of the coformer solubility.

Since cocrystal solubility is dependent to a large extent on the concentration of cocrystal components in the solution phase, especially for highly soluble cocrystals (relative to the drug), assessing the solubility product K_{sp} is more reliable for cocrystal solubility estimation. High-solubility cocrystals compared to the parent drug can significantly impact equilibrium concentrations and may undergo solution (aqueous) mediated transformation to the less soluble drug, making solubility measurement a challenge.[109] The true cocrystal solubility is underestimated when transformation occurs during a solubility measurement in which the cocrystal is suspended in an aqueous solution. The drug concentrations observed during these kinetic measurements is highly dependent on the solution conditions and cannot be extrapolated to other solution conditions. Good et al.[110] developed a method for the determination of cocrystal solubility under equilibrium conditions. This is based on developing phase solubility diagrams and measuring the solution composition at the cocrystal transition or eutectic point. Shown in Fig. 4.4 is a plot of a drug against the coformer concentration. The cocrystal concentration decreases as a function of coformer concentration and intersects the drug concentration solubility curve at a point called the eutectic point, where the two solid phases (cocrystal and drug or coformer and the solution phase) are all in equilibrium at a constant temperature and pH. The eutectic point can be experimentally determined and is a key parameter that establishes the regions of

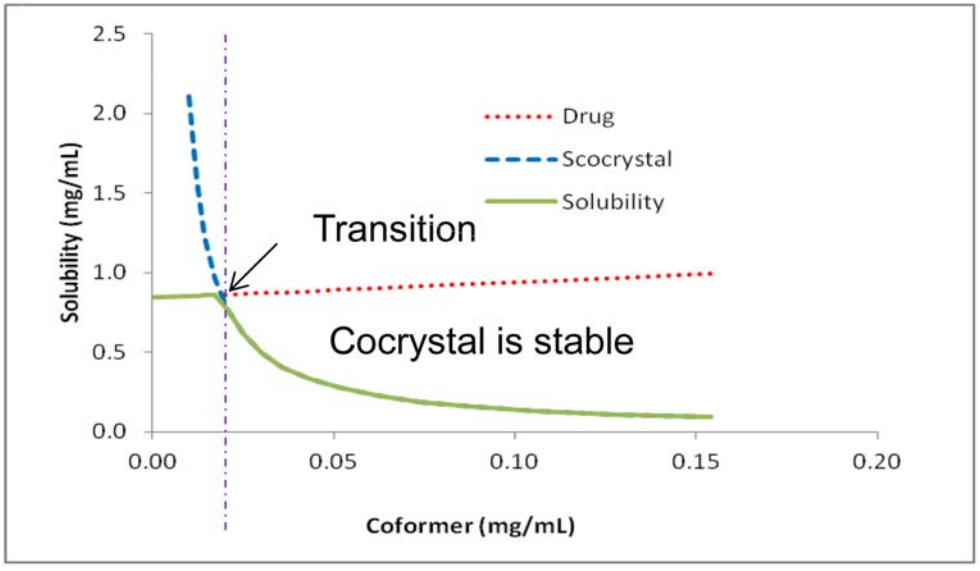


FIGURE 4.4 Phase solubility curve as a function of coformer concentration for the formation of cocrystals showing the transition or eutectic point.

thermodynamic stability of a cocrystal relative to its components. Similar to the pH_{max} in the case of salts, the eutectic point for cocrystals is a point where all phases, the solution, the solid cocrystal, and a solid component (drug or coformer), exist in equilibrium. This cocrystal eutectic point is an isothermal invariant point base on the Gibbs phase rule.

Multiple eutectic points can exist depending on what solid phases coexist at equilibrium. From the concentrations of drug and coformer at the eutectic point, the cocrystal solubility product K_{sp} can be calculated.

Cocrystal eutectic constant (K_{eu}), is the ratio of solution concentrations of cocrystal components at the eutectic point and is shown to be a function of the solubility ratio of cocrystal and drug in pure solvent. Similar to cocrystal solubility, it is shown to depend on solution chemistry including solvent, complexation, and ionization.[110] For a 1:1 cocrystal:

$$K_{eu} = \frac{[B]_{eu}}{[A]_{eu}}$$

Typically cocrystals are formed with acidic, basic, amphoteric, or zwitterionic coformers. Examples of coformers include but are not limited to benzoic, salicylic, maleic, glutaric, oxalic, fumaric, malic, and caprylic acids, to name a few.[111–113] Given the wide range of coformers similar to salts, cocrystal solubility can vary significantly based on ionization properties of coformers.

Cocrystals are typically made using solvent-based crystallization, mechanical grinding, or melt crystallization techniques. In solvent-based crystallization methods, it is predominantly recommended to utilize nonstoichiometric ratios of the components to create a large thermodynamic driving force for cocrystal formation. This use of nonstoichiometric composition has been termed a reaction crystallization method. Alternatively, stoichiometric ratios of reactants could be dissolved in a solvent of choice to achieve supersaturation either through a temperature difference or through evaporation of the solvent (solvothermal or evaporative methods). Further detail on screening strategies and cocrystal engineering is not covered here but can be found in a number of references.[8,32,111–115] Pharmaceutical cocrystals are becoming increasingly important as an alternate way to improve bioavailability and/or potential stability, mainly for weakly ionizable or nonionizable molecules with multiple ongoing efforts in supramolecular or crystal engineering to explore this.

4.2.2.3 *Polymorphs, solvates, and hydrates*

All crystalline forms such as neutral molecules, salts, or cocrystals can exist as polymorphs,[116] having the same chemical composition but different internal structures (lattice structures and conformations).[46] Polymorphs have been characterized based on structural differences in the crystal lattice as packing polymorphism or conformational polymorphism.[47] In the case of packing polymorphism, rigid molecules assemble into three-dimensional structures through different intermolecular mechanisms,[117] while in the other, flexible molecules are packed in different arrangements and are packed differently into the crystal structure.[118] When temperature is introduced as criteria, polymorphs can be classified as monotropic where the same polymorph is stable regardless of temperature, whereas for enantiotropic systems, there exists a transition temperature where the stability relationship inverts.

Solvates and hydrates are referred to as pseudopolymorphs where the crystalline form includes a solvent molecule such as water (hydrate) or any other solvent (solvate) in the crystal lattice. From a US regulatory perspective, solvates, hydrates, and polymorphs belong to the class of polymorphs.[119] Based on the structure crystalline hydrates can be categorized as channel hydrates, isolated-site hydrates or ion-associated hydrates.[47] Channel hydrates can be nonstoichiometric where additional water molecules can enter the channels when exposed to high humidity and dehydrate under low humidity, expanding or contracting the crystal lattice. Nonstoichiometric hydrates are of specific concern, if chosen for development, due to changes in water content in the crystals.[120]

For polymorphs, the free energy temperature diagrams and the Burger–Ramberger thermodynamic rules[121] are common to assess stability relationships. The relative thermodynamic stability of the forms at a constant temperature and pressure is defined based on the difference in the Gibbs free energy given as

$$\Delta G = \Delta H - T\Delta S$$

where ΔH is the enthalpy difference between forms as it relates to the lattice energy difference and ΔS relates to entropy difference or crystal disorder. A negative free energy or decrease in the free energy indicates a spontaneous process and is thermodynamically favored, while a positive value indicates a thermodynamically unfavored process. When $\Delta G = 0$, the system is at equilibrium, and the free energy of the two phases is the same.

Fig. 4.5 shows a relationship of the Gibbs free energy versus temperature for a solid that can exist in multiple crystalline and amorphous states, showing the relative thermodynamic stability of one related to the other.[116] The intersection points represent crystalline solids in equilibrium, with liquid denoted by T_m or transition temperatures of crystalline-to-crystalline or crystalline-to-amorphous transitions.[116]

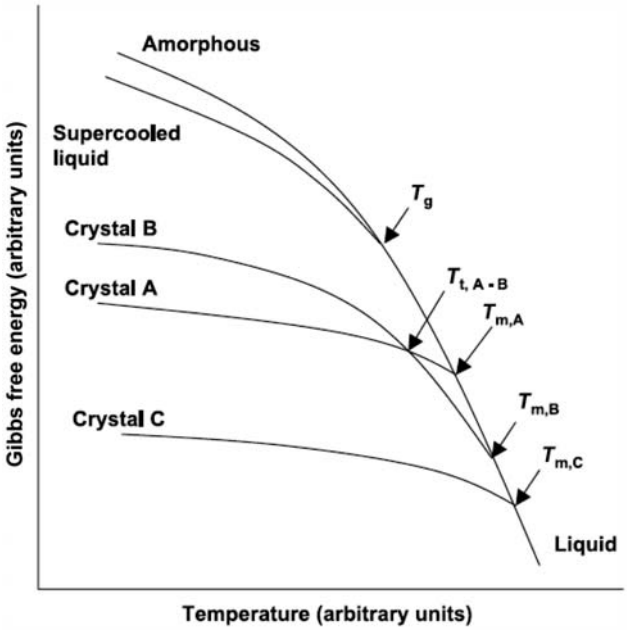


FIGURE 4.5 Schematic of Gibbs free energy curves for a hypothetical single-component system that exhibits crystalline and amorphous phase transitions. Monotropic systems (A and C, A and B), enantiotropic system (A and B) with transition temperature T_t, and an amorphous and supercooled liquid with a glass transition temperature T_g. Melting points, T_m, for the crystalline phases are shown by the intersection of the curves from the crystalline and liquid states. *Reprinted with permission Rodriguez-Spong B, et al. General principles of pharmaceutical solid polymorphism: a supramolecular perspective. Adv Drug Deliv Rev 2004;56:241–74.*

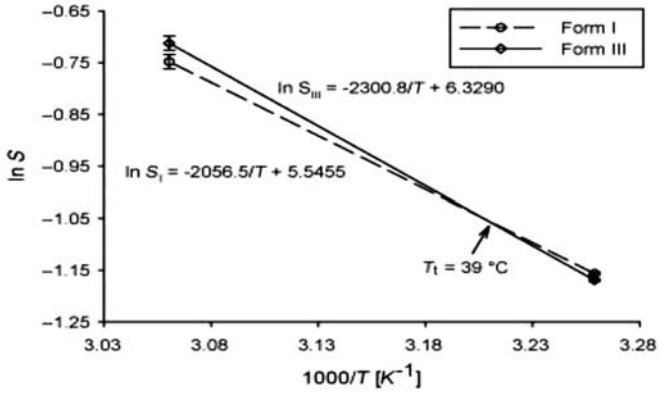


FIGURE 4.6 Van't Hoff plot showing monotropic and enantiotropic systems for form 1 and form II of flufenamic acid.[123] Reprinted with permission. *from Hu Y, et al. Estimation of the transition temperature for an enantiotropic polymorphic system from the transformation kinetics monitored using Raman spectroscopy. J Pharm Biomed Anal 2007;45(4):54651.*

For polymorphs, the thermodynamics of conversion or the stability of the various polymorphs is governed by ΔG, the Gibbs free energy. For crystalline solids that exist as multiple polymorphs, the polymorph with the lowest Gibbs frees energy will be the stable one under a specific set of conditions. The most stable polymorph will have the highest crystal lattice energy and the lowest solubility. All other intermediate solubility forms will be metastable and may be kinetically favored, but the thermodynamic stable form is still the one with the least solubility.

Depending on the relative dependence of each polymorph on temperature, it could be enantiotropic or monotropic. For a G versus T plot for two polymorphs, the monotropic forms always exist with one having a lower free energy than the other until the melting point, while for the enantiotropic forms, at some temperature on the G versus T curve, the stability relationship switches. At the point of intersection, the two forms are equally stable, and the $\Delta G = 0$. Van't Hoff plots[122] constructed using solubility or melting point data provide information on the stability of polymorphs, distinguishing between monotropic versus ennatiotropic systems. Fig. 4.6 shows the Van't Hoff plot for forms I and II for flufenamic acid with the transition temperature at 39°C.[123] From a practical consideration, during assessment of Van't Hoff plots using solubility measurements at various temperatures (log solubility vs $1/T$), care must be taken in the choice of solvents and the potential formation of solvates during solubility assessments. Also the phase purity of two polymorphs being evaluated needs to be known as well as any changes in the excess solid in the solubility measurements to ensure meaningful relationships.

Another method reported by Gu et al. is based on measuring the heat of a solution and solubility or dissolution rates at any one temperature[124] to estimate the transition temperature and stability between polymorphs. They demonstrated good agreement for the transition temperature of sulfamerazine polymorphs I and II to be between 51°C and 54°C determined by this method and confirmed experimentally. The choice of the method depends on the applicability to the system; for example, for thermally labile molecules and solvates, the melting point method is not suitable. Similarly, the method using solubility at various temperatures is unsuitable for polymorphs that undergo rapid transition in solution. All these methods are based on the assumption that the heat capacity difference between these is negligible.[122,124]

Solvates and hydrates can also have significantly different solubility and dissolution rates compared to each other and other neat forms. Incorporation of water or solvent molecules into the crystal lattice results in additional bonding with API molecules, altering crystal lattice density, internal energy and enthalpy, changing free energy, and the chemical potential of the solid. This may result in changes in physicochemical properties. Hydrate formation mainly depends on temperature and water activity. In

mixed-solvent systems, the activity of water is the key contributor for determining the phase transformation between hydrates and anhydrates. At a given temperature, at a given activity of water, solubility curves of the hydrate and anhydrates cross, and this is referred to as the critical water activity. For hydrates and anhydrates, to evaluate the risk of conversion, it is important to assess the critical water activity, which is the range of water activity where an anhydrate form switches to the hydrate form or vice versa. For a given drug at a given temperature, above the critical water activity (RH), a hydrate form will be more stable thermodynamically. The critical water activity can be determined by measuring the solubility of the anhydrate and the hydrate in a mixed-solvent system where accurate water activity measurements can be obtained and the solid phase monitored. Care needs to be taken that the mixed-solvent system selected does not have a propensity to form solvates. These studies may be long and time-consuming, and care must be taken to ensure that undesirable solvates are not formed during these experiments.

An alternative method by Li et al. proposed the use of this equation to estimate the critical water activity that requires determination of solubility ratios of hydrate and anhydrate[125]:

$$a_T[H_2O] = \sqrt[m]{\frac{x[A]}{x'[A]}}a[H_2O]$$

where, x[A] and x'[A] are the solubility of hydrate and anhydrate, m is the number of moles of water, $a_T[H_2O]$ is the activity of water at the transition point, and a [H_2O] is the activity of water at a given composition. Above the critical water activity for hydrate formation, hydrates typically are thermodynamically more stable and have lower solubility.[126,127] In an aqueous system, typically the solubility decreases as the order of hydrates increase, that is, an anhydrate form will have higher solubility than the monohydrate followed by its dihydrate. However, in a few cases, the hydrate may be intrinsically more soluble[128,129] than the anhydrate.

Estimates of the critical activity for systems that form hydrates and neat phases are important in the process design/control and in form selection.

4.2.2.4 *Amorphous forms*

During the form selection process, though crystalline forms are typically desired, some API solids may not crystallize under multiple conditions evaluated and only form amorphous solids. In other cases, none of the crystalline forms identified provide the necessary physicochemical properties such as solubility that are adequate to provide the desired in vivo exposures. In such cases where crystalline solid forms are not achievable or where they do not provide desired properties, an amorphous state may be considered.

Amorphous solids are defined as those that lack any crystalline structure and are composed of only disordered regions. They are known to have short-range molecular arrangements but lack long-range order. Amorphous forms possess excess thermodynamic properties such as enthalpy, entropy, and Gibbs free energy compared to the crystalline forms. In addition to the excess thermodynamic properties, amorphous forms also have greater mobility. Due to their high-energy state, amorphous forms could lead to increased apparent solubility, dissolution, and absorption. However, the physical stability of amorphous solids is a risk that requires characterization and mitigation. One common way to improve the physical stability is to make a solid dispersion with a polymer where the drug is dispersed in the polymer matrix.

An important parameter to assess the stability of amorphous systems is the glass transition temperature or T_g, which is the temperature below which the material is glassy or does not have molecular mobility and above which the amorphous solid is said to be in a rubbery state. Given that amorphous systems are high-energy systems, they possess higher solubility and dissolution rates compared to their crystalline forms and can provide a kinetic advantage when considering improved physicochemical properties of low-solubility drugs, thus providing higher in-vivo exposures. However, their physical stability needs to be considered, and maintaining the systems below their T_g ensures the material will be amorphous throughout the shelf life of the DP. Conversion of amorphous to crystalline is often known to be promoted by the presence of water, which can act as a plasticizer, lowering the T_g. This is another important consideration when evaluating the solid-state stability of amorphous systems when exposed to high humidity, since amorphous systems can adsorb a large amount of water, which can change the bulk properties. Since amorphous systems can adsorb large amounts of water, the chemical stability of these also needs evaluation, since larger amounts of water can be available for chemical reactions that are mediated with water.

Various processes are used to make amorphous material, including quenching of melts, rapid precipitation by antisolvent addition, freeze-drying, spray-drying, introduction of impurities, and solid dispersions. Mechanical or chemical process stresses such as grinding, milling, drying of crystalline hydrates, or wet granulation can also lead to amorphous formation.[130] Typical analytical methods used to assess amorphous systems are XRD, differential scanning calorimetry (DSC), solid state NMR, and modulated DSC, to name a few.[130]

It is well known in literature that polymers can potentially inhibit crystallization of the amorphous systems by either increasing T_g or by improved ionic or hydrogen bonding as well as potentially impacting the kinetics of crystallization.[131] Typically, it is assumed that for polymeric amorphous dispersions the ratio of T_g/T_m is around 0.75 or 0.8, and the dispersions are considered to be potentially stable against crystallization. T_g is a kinetic parameter. However, another important criteria to consider for stability of polymeric dispersions containing an amorphous drug is the thermodynamic driving force for crystallization and the activation energy for nucleation, which relates to the tendency of crystallization of the pure drugs themselves.[41,131,132] Sometimes it is difficult to crystallize, and an amorphous solid may be chosen as a form in this case. A detailed discussion on the selection, risks, and mitigation are outside the scope of this chapter.

4.3 SCREENING SOLID-FORMS OF API

Screening includes a wide range of API forms (polymorphs, hydrates, solvates, cocrystals, and salts) to ensure a reproducible, stable form with the desired physicochemical, biopharmaceutical, and powder properties can be obtained as part of this process. Typically, a thermodynamically stable form is the preferred choice, though in occasional cases, one may select a metastable form for improved pharmaceutical properties.

Balancing various properties such as crystallinity, hygroscopicity, stability and dissolution rate/solubility is needed to find an optimal form. Either a tiered or integrated approach is used toward screening. An integrated approach looks at a comparison of all key attributes across multiple forms, including crystallinity, polyphorphism, chemical and physical stability, hygroscopicity, and manufacturability and selecting an optimal form based on all these properties across multiple forms. While this approach is robust with polymorphism built into the early screen, it is extremely resource and material intensive. Fig. 4.7 shows an integrated approach to form selection.

In contrast to this, a tiered approach includes screening and eliminating various forms based on first assessment of crystallinity and solubility, followed by other properties such as thermal properties, hygroscopicity, stability scale-up feasibility, and bioavailability and finally leading to polymorph screening only on the lead candidates from the earlier tiered screens. This approach, while resource and material sparing, can result in eliminating some potentially useful forms early without complete characterization. Striking a balance between these approaches depending on the intrinsic properties of the molecule, key challenges, material availability, phase of development, and any

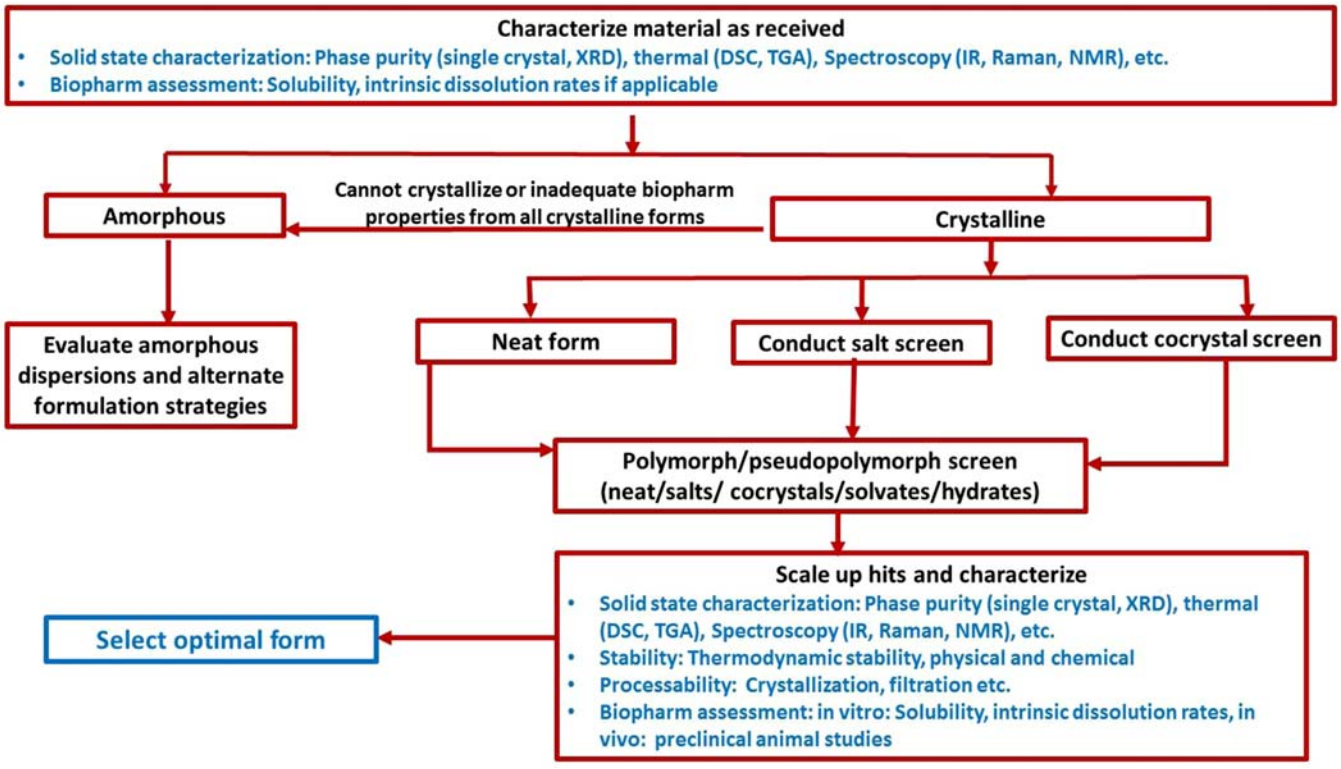


FIGURE 4.7 Integrated strategy for form screening and selection.

therapeutic area constraints as well as any formulation considerations is important during form screening.

4.3.1 Screening techniques

Crystallization is the common technique used to screen forms and is typically described by a nucleation step followed by crystal growth. For a polymorphic system, the polymorph that nucleates first has the fastest nucleation rate and the lowest free energy barrier to nucleation. However, the polymorph that actually crystallizes out depends on the combination of relative rates of nucleation and crystal growth. Hence, at a given time, multiple nuclei can exist, and if kinetic factors prevent equilibrium, a mixture of forms may be obtained. Various factors such as the degree of supersaturation during the process, temperature and cooling curves during the process, seeding, agitation, solvent composition, pH, polarity, and the presence of certain impurities or other additives that inhibit nucleation process have all shown to influence crystallization process and can be tailored to favor one form over another. The more stable form is obtained under thermodynamic conditions, while a metastable form is obtained under kinetic conditions. Typically, crystallization from solution (cooling or evaporative, slurry conversion) and recrystallization from neat compound (sublimation, thermal treatment, crystallization from melt, grinding, thermal, and desolvation) are common methods. The outcome of form from a crystallization experiment is influenced by both composition and process parameters, as shown in Table 4.6,[8] and can be tailored to obtain the thermodynamically stable or metastable forms. Supersaturation and temperature are the key driving forces that govern the kinetics of nucleation and crystal growth, thus influencing the form that crystallizes. Temperature, as discussed above, can generate different types of polymorphs for enantiotropic versus monotropic systems. The selection of the right solvent is also important and should cover a wide range of properties, considerations of solvents used in synthesis or formulation, and solvent water mixtures for assessment of potential hydrates. Another important factor that can play a role in solvent-based methods is the solubility of the API, and for highly soluble materials, an antisolvent addition may be a preferred way to crystallize. For salts and cocrystals, the solubility of the API and the counterion in the solvent both need to be considered while designing crystallization studies. To ensure better control on the crystallization process, understanding both thermodynamics and kinetics is important, and the conditions and parameters that control the crystallization process need to determined. Seeding with the desired form is often used to drive the process quickly toward the more stable polymorph.

Early screens are more qualitative in nature and may not translate into scalable processes, but they serve as a guide for future robust screens to pursue focused forms. These early experiments can also generate material for seeds for future manual assessments. Typical screens include high-throughput crystallizations, manual screens, or alternate screens.

4.3.2 High-throughput screening

Given the number of experimental parameters for form screening coupled with limited predictive capabilities, high-throughput crystallizations using an automated robotic setup offer an attractive option for a large array of forms and conditions to be explored in

TABLE 4.6 Crystallization Composition and Processing Variables[46,133]

Composition type		Process variables (applicable to all types of screens)				
Polymorphs & solvates	**Salts & cocrystals**	**Thermal**	**Antisolvent**	**Evaporation**	**Slurry conversion**	**Other variables**
• Solvent/solvent combinations • Degree of supersaturation • Additive type • Additive concentration	• Counterion type • Acid/base ratio • Solvent/solvent combinations • Degree of supersaturation • Additive type and concentration • pH (aqueous) • Ionic strength	• Heating rate • Cooling rate • Maximum temperature • Incubation temperature(s) • Incubation time	• Antisolvent type • Rate of antisolvent addition • Temperature of antisolvent addition • Time of antisolvent addition	• Rate of evaporation • Evaporation time • Carrier gas • Surface-volume ratio	• Solvent type • Incubation temperature • Incubation time • Thermal cycling and gradients	• Mixing rate • Impeller design • Crystallization vessel design (including capillaries, etc.)

Reprinted with permission Morissette SL, et al. High-throughput crystallization: polymorphs, salts, co-crystals and solvates of pharmaceutical solids. Adv Drug Deliv Rev 2004;56:275–300.

parallel. These experiments can cover a wide range of screens with limited material consumption given the small scale. Typically, high-throughput screening (HTS) screens are the first line of screens, and hits obtained from these studies can then be used to inform subsequent focused manual screens where material is further scaled up. Typical HTS include a polymorph screen, screening for acidic and basic counterions for salts (if applicable), cocrystal screen, and potentially even an amorphous dispersion screen if relevant. Screening is done in a 96-well plate in which solid (up to 10 mg) is dissolved in a suitable solvent, dispensed using the automated liquid handling systems. Supersaturation is varied by heating/cooling evaporation and slurry evaluations. Common methods such as thermal cooling crystallization, antisolvent, and evaporative crystallization are used. Thermal cooling crystallization includes the initial dissolution of API in the solvent followed by temperature ramps and controlled cooling to induce supersaturation and then crystallization. Antisolvent addition relies on the fact that an API is soluble to varying degrees in the crystallization medium but is largely insoluble in a particular solvent or solvents (eg, the antisolvent). Evaporative crystallization can be carried out by slowly increasing the concentration of API to achieve supersaturation and increase the degree of supersaturation (by preferential evaporation) in order to induce crystallization. In this method, differential rates of solvent loss from mixtures result in the unknown composition of the crystallization medium during nucleation, and the degree of supersaturation changes over the course of experiment, resulting in appearance of multiple crystal forms.[8] HTS is then coupled with optical image analysis and spectroscopic techniques such as Raman or XRPD that can read the high-throughput wells directly.[8] The initial characterization is followed by grouping of samples based on similar patterns for the wide number of crystallizations and provides a roadmap for further developing the desired form.

The advantage of these screens lies in the high throughput and wide range of the conditions analyzed. However, since these screens are typically meant to be fast screens and not always set under thermodynamic conditions, they can lead to metastable forms or cases with no crystalline hits. Considerations of form reproducibility, poor signal-to-noise ratio, and preferred orientation impact in spectra also need to be considered and sometimes limit the predictive capability on scale-up. There is literature supporting the use of proper design of experiments, and appropriate analysis for compounds with high-polymorphic states that HTS methods can demonstrate comparable results compared to the traditional intensive manual screens.[134]

4.3.3 Manual screens

Unlike the high-throughput automated screens, the manual screens are prepared as individual samples. They are conducted at a relatively larger scale than the HT screens and further identify the most stable or desirable form. A number of methods can be utilized such as solvent-based methods (most common) or solid-based methods such as subjecting the solid to various temperature or RH equilibration, grinding, compression or sublimation of solids, and also cooling from melt to assess the form that crystallizes out. Manual screening involves samples set up in individual vials or containers under a variety of conditions, including both solvent and solid.

Solvent-based methods are most widely used due to the large range of solvents and conditions that can be utilized. Solvent-based methods include cooling a solution, evaporation, antisolvent addition, vapor diffusion, and suspensions (slurries). Variations on these methods can include changes in solvent, solvent mixtures, antisolvent, temperature, cooling rate, concentration, rate of addition, and order of mixing, to name a few possibilities.[9] Nonsolvent methods consist of crystallization from melt, heat-induced transformations, sublimation, desolvation of solvates, salting out, and pH changes.

The goal of manual screens typically is to find the most stable form, and these studies include running under thermodynamic conditions such as slurrying for a few days, slow evaporation, slow cooling, slow antisolvent addition, or a combination of these methods. One of the common ways of ensuring the reaction reaches equilibrium with respect to the most stable form is seeding. Seeding is the addition of solids of the desired form into the crystallization to promote nucleation. For slurry conversion methods, solids are slurried and kept stirring in a wide range of solvents and aged for up to a few weeks, and the solid phase is assessed over time. A few limitations of the slurry method include the limitation of finding a stable form only at the experimental temperature, the presence of impurities that could prevent the stable form to crystallize, or very low solubility of API in solvents that could prevent transformations. Monitoring both the wet and dry cakes from slurry experiments ensures no transient solvates are missed, which dehydrate upon drying.

In the case of hydrates, exposure to RH, water slurry, temperature cycling of aqueous suspensions, mixed-solvent systems, and vapor diffusion by exposure of solid API to RH values ranging from 5% to 95% are commonly used to probe hydrate formation. For mixed-water-solvent systems, understanding the water activity is critical, since different solvent aqueous mixtures at a given mole fraction of water will have varying water

activity. (A 0.2 mol fraction of water in methanol gives 0.25 water activity, while in isopropanol, it will give 0.45.) This can lead to the formation of different stoichiometric hydrates or anhydrates at the same water content in various solvents.[9,135,136]

4.3.4 Alternate screens

Recently, several alternate techniques such as capillary crystallization, laser-induced crystallization, and sonocrystallization that promote nucleation have been explored to generate newer forms. More detailed discussion on these screens can be found in the references listed here.[9,116]

4.4 IDENTIFICATION AND ANALYSIS OF FORMS

Various analytical techniques used for assessing solid forms include a variety of physical, thermal, spectroscopic, and diffractometric methods. Typically, multiple techniques are used for complete characterization of the solid covering both molecular level (spectroscopic or single-crystal XRD) and particulate level (XRD, thermal methods, or microscopy). Thermal (DSC/TGA/modulated DSC), spectroscopic (NIR, IR, Raman, NMR), XRD, and microscopy-based techniques are the most common used. In addition, DVS to assess hygroscopicity is also an important characterization tool. Sometimes specialized techniques such as terahertz pulsed spectroscopy (TPS) or terahertz time-domain spectroscopy (THz-TDS), which is a vibrational spectroscopic technique, can be used to assess particulate-level, small structural changes. Techniques typically used for solid-state characterization are listed in Table 4.7.

Solubility/dissolution, stability, and hygroscopicity are key characterization elements that have already been described in detail in earlier sections. The next sections briefly describe commonly used analytical techniques during form selection process.

4.4.1 Single-crystal and PXRD

A single crystal is defined as a crystal consisting of a noninterrupted repetition of the unit cell in three dimensions. The single-crystal XRD method provides information on the structure of the unit cell, bond lengths and angles, molecular conformation, molecular packing, hydrogen bonding pattern, density, and crystal disorder. To acquire data from a single crystal, the sample is mounted, and a narrow beam of X-rays is passed through it and diffracts against the sample, and this diffracted beam is collected on the detector. Calculated powder-diffraction patterns from single-crystal data provide information on preferred

TABLE 4.7 Solid-State Characterization Methods

Method	Data measured	Measurement use
X-ray diffraction (single-crystal and PXRD)	Diffractogram	Crystallographic properties
Infrared (IR) spectroscopy	IR spectrum	Chemical information
Raman spectroscopy	Raman spectrum	Chemical information
Terahertz pulsed spectroscopy	Terahertz pulsed spectrum	Chemical information, lattice photon modes
Near-infrared (NIR) spectroscopy	NIR spectrum	Chemical information (overtones and combinations of IR vibrations)
Solid-state NMR	Magnetic resonance	Chemical information
Differential scanning calorimetry	Heat flow vs temperature	Thermal events
Thermogravimetry (TG)	Change of mass vs temperature	Solvate/hydrate studies
Microscopy, polarized light microscopy (PLM), scanning electron microscopy (SEM)	Microscopy under the influence of light or electron radiation	Morphology, surface examination, dehydration, polymorphism (PLM)
Microcalorimetry	Heat flow vs time	Quantification of amorphous form
Solution calorimetry	Heat flow during dissolution	Quantification of polymorphs and amorphous form
Moisture sorption/desorption isotherms	Change of mass vs RH%	Hygroscopicity, hydrate formation, dehydration, amorphous crystallization
Solubility/dissolution	Amount dissolved in different solvents or temperatures/time	Solubility or dissolution rate measurement

Reprinted with permission from Aaltonen J, et al. Solid form screening—a review. Eur J Pharm Biopharm 2009;71:23–37.

orientation and phase purity. *Phase purity* refers to the presence of a single dominant crystalline phase, which is a desirable attribute for an API. Phase impurities are typically low levels of other forms that can be present along with the desired form and impact the stability and potentially the performance of the DP. Single-crystal X-ray determination offers a powerful tool to assess the phase purity.

Determining the crystal structure using this approach is the most common and accurate way of understanding the three-dimensional structure, and it is important to understand the physical properties of one crystalline form versus another. It also shows the intermolecular interactions of the atoms within the crystal. The XRD pattern obtained using the single-structure analysis serves as a reference diffractogram The technique does rely on obtaining high-quality crystals of adequate size and hence needs crystals grown to be able to get crystal structure.

PXRD is one of the most common bulk techniques used for the analysis of powders and is similar in concept to the single-crystal analysis, except the data is generated on the bulk powder instead of a single crystal. The diffractogram obtained for a given form contains information about the unit cell and is extremely sensitive to changes in the cell parameters, which makes it a gold standard for screening polymorphs during form selection and the drug development process. Crystalline material shows sharp, intense peaks, while amorphous material shows a broad peak or peaks, called an amorphous halo. This also makes PXRD a technique of choice to understanding the crystallinity of a given sample.

4.4.2 Thermal techniques

The most common and routinely used thermal techniques are DSC and thermogravimetric analysis (TGA). The techniques are based on heating the solids and monitoring changes in the physical properties. In DSC the difference in temperature between the sample and the reference as the temperature is changed is monitored to provide information on enthalpy change. TGA provides information on the change in mass of a solid as a function of temperature.

The DSC technique measures heat absorbed or released on heating, cooling, or maintaining a sample at a constant temperature. Thermodynamic parameters such as melting point, heat capacity, and heat of fusion can be obtained by this technique in addition to transition events that occur due to phase transitions. The endotherm represents a process where heat is absorbed such as melting, solvent loss, or phase transitions, while exotherm represents a process where heat is liberated such as recrystallization or chemical reactions where heat is involved. The area under the peak is proportional to heat change and gives enthalpies of transition. In the case of amorphous solids, the glass transition temperature is seen on the DSC curve by a change in heat capacity.

The method can be used for accurate melting point and purity assessment where phase impurities can be determined by additional shifts in endothermic peaks. Additional experimental variables such as pan configuration (open or closed), heating rates, and sample and particle size can also affect the shape of DSC curves. For materials that are partially amorphous or mixtures of crystalline and amorphous, glass transition events could coincide with other endothermic or exothermic events, and modulated DSC which separates the thermodynamic and kinetic components, can be used. In recent years, hyper-DSC or high-speed DSC has become popular given its ability to scan at extremely fast rates of 100−500°C/min range, which have enabled more sensitivity to some transitions.[20,137] DSC is a powerful tool to assess the relationship between polymorphs (enantiotropic vs monotropic), evaluate mixtures of polymorphs, physical stability of amorphous dispersions, and quantify amorphous content.[138,139] In some cases, traditional DSC scan rates are too slow and may cause some materials to recrystallize during the melting process or decompose immediately after melting. In such cases, the HyperDSC (fast-scan DSC) technique, which uses scanning rates as fast as 100−500°C/min in heating as well as cooling over a broad temperature range, has been used.[20,137,140,141] These fast scanning rates inhibit kinetically controlled processes such as recrystallization during melting and decomposition after melting, which are either completely eliminated or significantly reduced by this method. Thus, HyperDSC has been successfully used to detect small amorphous contents in crystalline material,[137] study thermal properties between polymorphs as in the case of carbamazepine,[140] or understand polymorphic phase impurities in crystalline material.[141]

Another most commonly used technique is the thermogravimetric analysis. TGA analysis that monitors change in mass with temperature can be run in the isothermal mode where the temperature is kept constant or nonisothermal (dynamic) mode where the temperature is raised, typically linearly at a fixed rate. Similar to the DSC, the TGA setup including heating pan configuration, sample and particle size, and heating rate can all affect the TGA curves. Typically a dry nitrogen purge is used during the study to minimize the impact of RH, atmospheric oxygen, or liberated gases to further impact the sample. TGA curves run in isothermal mode coupled with DSC curves can be used to provide additional information on

endothermic events such as melt versus decomposition or desolvation. Additionally, isothermal curves of weight loss as a function of time can be used to determine reaction rates.

Another thermal technique is microcalorimetry, which measures heat liberated or taken up during various physical or chemical processes. For example, heats of solution estimation to determine the thermodynamic stability between polymorphs, amorphous-to-crystalline transitions, or quantification of amorphous in crystalline material can all be detected by this technique since it is very sensitive in small changes in the heat taken up or liberated.[142] In isothermal microcalorimetry, the sample and the reference are both maintained at a constant temperature, and the heat released or absorbed by the progression of the reaction is quantitatively measured via thermocouples as a function of time.[21]

4.4.3 Spectroscopic techniques

Various spectroscopic techniques are routinely used to probe the solid state forms and short-range structure in molecular solids. Raman and IR spectroscopy probe the solid state predominantly on the intramolecular level. XRD and the various thermal analytical techniques predominantly probe the intermolecular level.[23]

Raman and IR spectroscopy are the most common vibration spectroscopic techniques, and they provide information on the structure and the molecular conformation in the solid state. In case of polymorphs, these techniques are useful in distinguishing the differences in hydrogen bonding or crystal lattice energies based on the differences in packing between various polymorphs. Near infrared (NIR) is one of the most advanced spectroscopic techniques used as a PAT tool. Raman spectroscopy can be used to monitor crystallization process. Solid-state nuclear magnetic resonance (SSNMR) is another powerful technique that provides short-range information about the solid-state structure primarily. Various nuclear isotopes are used such as ^{13}C, ^{1}H, ^{15}N, ^{19}F, ^{23}Na, and ^{31}P. SSNMR spectroscopy has been applied in a number of pharmaceutical applications including studies of polymorphism, understanding phase purity, amorphous content, solvation, salts, cocrystals, and various types of formulations.[143]

Initial assessments include single-crystal analysis to determine crystalline structure, obtain phase purity information, and understand polymorphism for a given API solid. Once this has been determined, additional spectroscopic techniques such as Fourier transform infrared spectroscopy (FTIR), Raman, and SSNMR can be employed to further characterize the various solids. Alongside thermal methods such as DSC and TGA, ITC can provide insights into assessing physical stability as in case of salts and cocrystals determine transitions between metastable and stable forms, and assess thermodynamic stability including enantiotropic/monotropic relationships between polymorphs. As described above, a combination of these analytical techniques are used during the form selection process to completely characterize these forms.

4.5 CONCLUSIONS

Drug substance solid forms should be selected not only to enable drug substance quality and manufacturability but also DP considerations. Ideally, the selected form possesses optimal physical, chemical, mechanical, and biopharmaceutical properties. The formation of salts, cocrystals, solvates, and hydrates as well as existence of polymorphs is rarely predictable. Therefore, the selection process typically involves broad screening for solid forms and characterization including thermodynamics and kinetics as well performance assessment with a view toward selecting a solid form that is not only suitable as a drug substance but also an optimal DP dosage form. Screening may be manual or automated in a high-throughput fashion and utilize a variety of techniques such as crystallization, precipitation, and thermal techniques. The screening may be broad; however, the experimental design is based on some physical-chemical properties in choosing solvents and concentrations as well as counterions and coformers. Several solid-state techniques and solid/solution experimental approaches are used to assess thermodynamics and kinetic stability. Developing an understanding of thermodynamics of polymorphs, hydrates, and solvates (in process-relevant solvents) is important to completely avoid or minimize risks. In addition to these characterization studies, the quality and performance of a solid form is assessed by conducting accelerated stability testing, solubility/dissolution testing, and, if necessary, in-silico and in vivo studies for pharmacokinetics evaluation. A thorough solid-form screening, characterization, and selection are essential to maximize the potential of enabling a drug candidate from patient, quality, and manufacturing perspectives.

4.6 CASE STUDIES

4.6.1 Case study 1: RPR111423[144]

RPR 111423 is a very weak base with a pK_a of 4.25. The strategy was to first carry out a comprehensive screening of various counterions and solvent matrix to

identify two crystalline salts: monohydrochloride and a mesylate. The properties of these salts and the free base were then studied and compared.

The results showed that both salts significantly enhance solubility but exist in four to six different polymorphs. They both readily convert to free base in water and intestinal fluids. The corresponding results for the free base indicated that it appeared to be the better candidate. It showed no evidence of polymorphism, and it was not hygroscopic. A subsequent study performed on samples of the drug substance and on simple capsule formulations demonstrated that the dissolution rates of the micronized free base were equivalent or superior to those of the salts under the same conditions. Thus, the free base was selected for further development.

A few key observations can be drawn from this case study. First of all, identifying what properties to modify by forming a salt prior to any salt screening can save time and resources; a salt may not always be needed. In-vitro dissolution experiments and in-silico simulation can be very useful tools to evaluate the need for a salt from a PK point of view. Additionally, a pharmacokinetic study in animals comparing a solution formulation to a suspension of the micronized free base at the projected clinical dose can provide further indication of whether a salt is needed. However, if the projected dose gets much higher for this compound, a salt would provide advantages. Sensitivity analysis using GastroPlus software or the estimation of the maximum absorbable dose can be good tools in understanding at what doses solubility and dissolution may become rate-limiting to absorption.

4.6.2 Case study 2: LY333531[145]

The following case study is extracted from an excellent article by Engel et al., which reported an integrated, tiered-approach to select a salt form of a drug candidate.

LY333531 is a weakly basic drug with low solubility of the free base ($<1\ \mu g/mL$). Although the pK_a of the drug is not reported by the authors, it is estimated to be 9−11, based on the presence of a tertiary amine. Based on the need for oral dosage form of this BCS class II drug, the authors evaluated salt forms. They crystallized small lots (<1 g) of hydrochloride, sulfate, mesylate, succinate, tartrate, acetate, and phosphate salts. Preliminary characterization using polarizing microscopy, thermal analysis, PXRD and solubility was conducted on all seven salts. Based on the poor crystallinity, low solubility, and difficulty in chemical purification, only two salts (namely, hydrochloride and mesylate) were chosen for further consideration. The two salts, hydrochloride and mesylate, were subjected to further characterization including polymorphism, hydrate formation and hygroscopicity, stability, purification, filterability, and pharmacokinetics in dogs. Although methanol/water was initially used to prepare the mesylate salt, an alternate-solvent system using acetone/water (9:1 v/v) was chosen to avoid generation of methyl methanesulfonate, a mutagen. Three forms of hydrochloride forms with varying degrees of hydration were identified: an anhydrate form and two hydrates. The two interconverting hydrates were tentatively assigned to be monohydrate and tetrahydrate based on the vapor pressure isotherm data, although the exact structures of these hydrates were not determined. The anhydrate form did not show significant hygroscopicity (water uptake) during the vapor isotherm experiments and was considered for further stability and bioavailability assessments. Only a monohydrate form of mesylate salt was reported, and it showed no significant water uptake up to 90% RH. No change in PXRD pattern was seen at the end of the vapor isotherm experiments, although a small amount of hysteresis was seen during the desorption phase. The aqueous solubility of mesylate (0.5 mg/mL) was 5X higher than the hydrochloride salt (0.1 mg/mL). The chemical stability of the two salts as well as blends with three excipients was found to be acceptable after 1 month of storage at 40°C/75% RH and 50°C. Preliminary crystallization and filtration studies on both salts indicated that while there was no filtration advantage of one salt over the other, the purification was a lot better with the mesylate salt than the hydrochloride salt. The mesylate salt was found to be 2.6X more bioavailable compared to the hydrochloride salt when dosed as an oral suspension at 20 mg/kg in dogs. The mesylate salt monohydrate form was chosen as the API for further development.

4.6.3 Case study 3[146]

The following case study is summarized based on the work of McNamara et al. The authors demonstrated the screening, characterization, and evaluation of glutaric acid cocrystal to improve oral bioavailability of a low solubility API.

Compound 1 is a poorly soluble drug with $<0.1\ \mu g/mL$ in aqueous media. Salt formation was not possible even with strong acids such as hydrochloric, sulfuric, or phosphoric acids because of the low pK_a (estimated to be -0.7). The authors made some attempts to make amorphous forms but decided to look at cocrystals because of the risk of crystallization of the amorphous solid (low T_g). The authors screened for cocrystals with 26 different carboxylic acid-based coformers using the binary melt method on a hot-stage microscope. Five different new solid phases

(cocrystals) were obtained and analyzed by spectroscopy. The authors selected the glutaric acid cocrystal[2] for further evaluation because of the relatively high melting point of the cocrystal and the expected high water solubility of the cocrystal because of the high water solubility of the coformer. A solvent-based crystallization was developed to scale-up the process to make gram quantities of cocrystal.

Single-crystal X-ray data showed that one carboxylic group of the coformer hydrogen bonds to the amide functionality in an eight-membered ring motif, and the second acid group forms an interaction with a pyrimidine acceptor site. DSC thermal scans revealed that the cocrystal[2] melts at 142°C, whereas compound 1 melts at 206°C. The cocrystal[2] showed a significantly higher dissolution rate in water compared to compound 1. The intrinsic dissolution of the cocrystal was 18 times greater than compound 1.

The cocrystal[2] was found to be nonhygroscopic as it sorbed less than 0.08% even at 95% RH. Both chemical and physical stability was good after 2 months of storage at 40°C/75% RH and 60°C. Though the cocrystal[2] did convert to compound 1 in 37°C liquid water after 24 hours, the cocrystal[2] did not convert to compound 1 during the 90 minutes required for the dissolution test or after 6 hours of exposure to 100% water vapor at RT. A dog PK study showed a significant increase in plasma concentration values after oral dosing of 2 relative to 1 at doses of 5 and 50 mg/kg. In summary, the authors reported the viability of the cocrystal as a crystalline solid form of a poorly soluble drug that was not amenable to salt formation.

References

1. Aguiar AJ, Zelmer JE. Dissolution behavior of polymorphs of chloramphenicol palmitate and mefenamic acid. *J Pharm Sci* 1969;**58**(8):983–7.
2. Bartolomei M, et al. Physico-chemical characterisation and intrinsic dissolution studies of a new hydrate form of diclofenac sodium: comparison to the anhydrous form. *J Pharm Biomed Anal* 2006;**40**(5):1105–13.
3. Badawy S. Effect of salt form on chemical stability of an ester prodrug of a glycoprotein IIb/IIIa receptor antagonist in solid dosage forms. *Int J Pharm* 2001;**223**(1–2):81–7.
4. Childs SL, Kandi P, Lingireddy SR. Formulation of a danazol cocrystal with controlled supersaturation plays an essential role in improving bioavailability. *Mol Pharm* 2013;**10**:3112–27.
5. Kobayashi Y, et al. Physicochemical properties and bioavailability of carbamazepine polymorphs and dihydrate. *Int J Pharm* 2000;**193**:137–46.
6. Suihko E, et al. Dynamic solid state and tableting of four theophylline forms. *Int J Pharm* 2001;**217**:225–36.
7. Aaltonen J, et al. Solid form screening—a review. *Eur J Pharm Biopharm* 2009;**71**:23–37.
8. Morissette SL, et al. High-throughput crystallization: polymorphs, salts, co-crystals and solvates of pharmaceutical solids. *Adv Drug Deliv Rev* 2004;**56**:275–300.
9. Newman A. Specialized solid form screening techniques. *Org Process Res Dev* 2013;**17**:457–71.
10. Pfannkuch F. Biological effects of the API salt form. *Handbook of pharmaceutical salts* 2011;125–46.
11. Stahl P, Nakano M. Pharmaceutical aspects of the drug salt form. In: Stahl P, Wermuth CG, editors. *Handbook of Pharmaceutical salts*. Weinheim: Weily -VCH; 2006. p. 84–116.
12. Seajuddin ATM, Pudipeddi M. Salt selection strategies. In: Stahl P, Wermuth CG, editors. *Handbook of Pharmaceutical Salts*. Weinheim: Wiley-VCH; 2002. p. 135–60.
13. Byrn SR, Xu W, Newman AW. Chemical reactivity in solid-state pharmaceuticals: formulation implications. *Adv Drug Deliv Rev* 2001;**48**:115–36.
14. Childs SL, Stahly GP, Park A. The salt-cocrystal continuum: the influence of crystal structure on ionization state. *Mol Pharm* 2007;**4**(3):323–38.
15. Elder DP, Holm R, Diego HL. Use of pharmaceutical salts and cocrystals to address the issue of poor solubility. *Int J Pharm* 2013;**453**:88–100.
16. Gould PL. Salt selection for basic drugs. *Int J Pharm* 1986;**33**:201–17.
17. Serajuddin ATM. Salt formation to improve drug solubility. *Adv Drug Deliv Rev* 2007;**59**:603–16.
18. Shan N, Zawortko J. The role of cocrystals in pharmaceutical science. *Drug Discov Today* 2008;**13**(9-10).
19. Yu L, Reutzel SM, Stephenson GA. Physical characterization of polymorphic drugs: an integrated characterization strategy. *Pharm Sci Technol Today* 1998;118–27.
20. Ford JL, Mann TE. Fast-Scan DSC and its role in pharmaceutical physical form characterization and selection. *Adv Drug Deliv Rev* 2012;**64**:422–30.
21. Phipps MA, Mackin LA. Application of isothermal microcalorimetry in solid state development. *Pharm Sci Technol Today* 2000;**3**(1):9–17.
22. Vogt FG, et al. New approaches to the characterization of drug candidates by solid-state NMR. In: Gad SC, editor. *Pharmaceutical sciences encyclopedia: drug discovery, development, and manufacturing*. New York, NY: John Wiley & Sons, Inc.; 2011. p. 1–51.
23. Aaltonena J, et al. Perspectives in the use of spectroscopy to characterise pharmaceutical solids. *Int J Pharm* 2008;**364**:159–69.
24. Sun C, Grant DJW. Influence of crystal structure on the tableting properties of sulfamerazine polymorphs. *Pharm Res* 2001;**18**:274–80.
25. Ragnarsson G, Sjoegren J. Compressibility and tablet properties of two polymorphs of metoprolol tartrate. *Acta Pharm Suec* 1984;**21**(6):321–30.
26. Beyer T, Day GM, Price SL. The prediction, morphology, and mechanical properties of the polymorphs of paracetamol. *J Am Chem Soc* 2001;**123**:5086–94.
27. Di Martino P, et al. A new pure paracetamol for direct compression: the orthorhombic form. *Int J Pharm* 1996;**128**:1–8.
28. Ward GH, Schultz RK. Process induced crystallinity changes in albuterol and its effect on powder physical stability. *Pharm Res* 1995;**12**:773–9.
29. Qiu Y, et al. Effect of physicochemical and formulation variables on the in vivo absorption of ABT-761. *Drug Dev Ind Pharm* 2002;**28**:287–95.
30. Hancock BC, Zografi G. Effects of solid-state processing on water vapor sorption by aspirin. *J Pharm Sci* 1996;**85**:246–8.
31. Wostheinrich K, Schmidt PC. Polymorphic changes of thiamine hydrochloride during granulation and tableting. *Drug Dev Ind Pharm* 2001;**27**:481–9.
32. Childs SL, et al. Screening strategies based on solubility and solution composition generate pharmaceutically acceptable cocrystals of carbamazepine. *Cryst Eng Comm* 2008;**10**:856–64.

33. Murphy D, et al. Solution-mediated phase transformation of anhydrous to dihydrate carbamazepine and the effect of lattice disorder. *Int J Pharm* 2002;**246**:121–34.
34. Law D, et al. Ritonavir–PEG 8000 amorphous solid dispersions: in vitro and in vivo evaluations. *J Pharm Sci* 2004;**93**(3):563–70.
35. Matsuda Y, et al. Pharmaceutical evaluation of carbamazepine modifications: comparative study for photostability of carbamazepine polymorphs by using Fourier-transformed reflection-absorption infrared spectroscopy and colorimetric measurement. *J Pharm Pharmacol* 1993;**46**:162–7.
36. Chen X, et al. Reactivity differences of indomethacin solid forms with ammonia gas. *J Am Chem Soc* 2002;**124**:15012–19.
37. De Villiers MM, Van der Watt JG, Lotter AP. Kinetic study of the solid-state photolytic degradation of two polymorphic forms of furosemide. *Int J Pharm* 1992;**88**(1–3):275–83.
38. Guo Y, Byrn S, Zografi G. Physical characteristics and chemical degradation of amorphous quinapril hydrochloride. *J Pharm Sci* 2000;**89**:128–43.
39. Xu W. *Investigation of solid-state stability of selected bioactive compounds. Ph.D. Thesis.* Purdue University; 1997.
40. Tong P, Zografi Z. Effects of water vapor absorption on the physical and chemical stability of amorphous sodium indomethacin. *AAPS PharmSciTech* 2003;**5**(2):9–16.
41. Ahlneck C, Zografi G. The molecular basis of moisture effects on the physical and chemical stability of drugs in solid state. *Int J Pharm* 1990;**62**:87–95.
42. Wu Y, et al. Reactive Impurities in Excipients: Profiling, Identification and Mitigation of Drug–Excipient Incompatibility. *AAPS Pharm Sci Tech* 2011;**12**(4):1248–63.
43. Nord K, Andersen H, Tonnesen H. Photoreactivity of biologically active compounds. XII. Photostability of polymorphic modifications of chloroquine diphosphate. *Drug Stability* 1997;**1**:243–8.
44. Byrn SR, et al. The crystal structure, solid-state, NMR spectra and oxygen reactivity of five crystal forms of prednisolone tert butyl acetate. *J Am Chem Soc* 1988;**110**:1609–14.
45. Visalakshi NA, et al. Behavior of moisture gain and equilibrium moisture contents (EMC) of various drug substances and correlation with compendial information on hygroscopicity and loss on drying. *Pharm Dev Technol* 2005;**10**(4):489–97.
46. Bryn SR, Pfeiffer RR, Stowell JG. s.l. *Solid state chemistry of drugs*. West Lafayette, IN: SSCI Inc; 1999.
47. Vippagunta SR, Brittain HG, Grant DWJ. Crystalline solids. *Adv Drug Deliv Rev* 2001;**48**:3–26.
48. Authelin J-R. Thermodynamics of non-stoichiometric pharmaceutical hydrates. *Int J Pharm* 2005;**303**:37–53.
49. Newman AW, Reutzel-Edens SM, Zografi G. Characterization of the "hygroscopic" properties of active pharmaceutical ingredients. *J Pharm Sci* 2008;**97**(3):1047–59.
50. Giron D, et al. Solid state characterizations of pharmaceutical hydrates. *J Therm Anal Calorim* 2002;**68**:453–65.
51. Reutzel-Edens SM, Newman AW. Physical characterization of hygroscopicity in pharmaceutical solids. In: Hilfiker R, editor. *Polymorphism in the Pharmaceutical Industry*. Weinheim: Wiley-VCH; 2006. p. 235–58.
52. Rohrs BR, et al. Tablet dissolution affected by a moisture mediated solid-state interaction between drug and disintegrant. *Pharm Res* 1999;**16**(12):1850–60.
53. Katdare AV, Bavitz JF. Hydrate related dissolution characteristics of Norfloxacin. *Drug Dev Ind Pharm* 1984;**10**(5):789–807.
54. Rahman Z, et al. Physico-mechanical and Stability Evaluation of Carbamazepine Cocrystal of nicotinamide. *AAPS Pharm Sci Tech* 2011;**12**(2):693–703.
55. Pudipeddi M, Serajuddin AT. Trends in solubility of polymorphs. *J Pharm Sci* 2005;**94**(5):929–39.
56. Brittain H, Grant DWJ. Effects of polymorphism and solid state solvation on solubility and dissolution rate. In: Brittain H, editor. *Polymorphism in pharmaceutical solids, drugs and the pharmaceutical sciences*, vol. 95. New York, NY: Marcel Dekker; 1999. p. 279–330.
57. Brittain H, Fiese EF. Effects of pharmaceutical processing on drug polymorphs and solvates. In: Brittain H, editor. *Polymorphism in pharmaceutical solids, drugs and the pharmaceutical sciences*, vol. 95. New York, NY: Marcel Dekker; 1999. p. 331–61.
58. Greco K, Bogner R. Solution-mediated phase transformation significance during dissolution and bioavailability. *J Pharm Sci* 2012;**101**(9):2996–3018.
59. Dezani AB, et al. Equilibrium solubility versus intrinsic dissolution: characterization of lamuvidine, stavudine and zidovudine for BCS classification. *Braz J Pharm Sci* 2013;**49**(4):853–63.
60. Kim M, et al. Preparation, characterization and in vivo evaluation of amorphous atorvastatin calcium nanoparticles using supercritical antisolvent (SAS) process. *Eur J Pharm Biopharm* 2008;**69**(2):454–65.
61. Klein S. The use of biorelevant dissolution media to forecast the in vivo performance of a drug. *AAPS J* 2010;**12**(3):397–406.
62. Galia E, et al. Evaluation of various dissolution media for predicting in vivo performance of class I and II drugs. *Pharm Res* 1998;**15**(5):698–705.
63. Bighley LD, Berge SM, Monkhouse DC. Salt forms of drug and absorption. In: Swarbrick J, Boylan JC, editors. *Encyclopedia of pharmaceutical technology*, vol. 13. New York, NY: Marcel Dekker; 1995. p. 453–99.
64. Chiang P, Wong H. Incorporation of physiologically based pharmacokinetic modeling in the evaluation of solubility requirements for the salt selection process: a case study using phenytoin. *AAPS J* 2013;**15**(4):1109–18.
65. Serajuddin ATM, Rosoff J. pH—solubility profile of papaverine hydrochloride and its relationship to the dissolution rate of sustained-release pellets. *J Pharm Sci* 1984;**73**(9):1203–8.
66. Hsieh Y-L, et al. Impact of sertraline salt form on the oxidative stability in powder blends. *Int J Pharm* 2014;**461**(1–2):322–30.
67. Serajuddin ATM, Jarwosky CI. Effect of diffusion layer pH and solubility on the dissolution rate of pharmaceutical acids and their hydrochloride salts I: phenazopyridine. *J Pharm Sci* 1985;**74**(2):142–7.
68. Serajjudin ATM, Jarowsky CI. Effect of diffusion layer pH and solubility on the dissolution rate of pharmaceutical acids and their sodium salts II: salicyclic acid theophylline and benzoic acid. *J Pharm Sci* 1985;**74**(2):148–54.
69. Kramer SF, Flynn GL. Solubility of organic hydrochlorides. *J Pharm Sci* 1972;**61**(12):1896–904.
70. Pudipeddi M, et al. Solubility and dissolution of weak acids, bases and saltss.l. In: Stahl PH, Wermuth CG, editors. *Handbook of pharmaceutical salts properties, selection and use*. Wiley-VCH; 2002.
71. Anderson BD, Conradi RA. Predictive relationships in the water solubility of salts of a non steroidal anti inflammatory drug. *J Pharm Sci* 1985;**74**(8):815–20.
72. Ledwidge MT, Corrigan OL. Effects of surface active characteristics and solid state forms on the pH solubility profiles of drug–salt systems. *Int J Pharm* 1998;**174**:187–200.
73. Black SN, et al. Structure, solubility, screening, and synthesis of molecular salts. *J Pharm Sci* 2007;**96**(5):1053–68.
74. Stahl P. *Preparation of water-soluble compounds through salt formation. The Practice of Medicinal Chemistry*. 2nd ed. Academic Press; 2003. p. 601–15s.l.
75. Tong W-Q, Whitesell G. In situ salt screening-a useful technique for preformulation and discovery support. *Pharm Dev Technol* 1998;**3**(2):215–23.

76. Cruz-Cabeza AJ. Acid–base crystalline complexes and the pK_a rule. *Cryst Eng Comm* 2012;**14**:6362–5.
77. Paulekuhn GS, Dressman JB, Saal C. Trends in active pharmaceutical ingredient salt selection based on analysis of the orange book database. *J Med Chem* 2007;**50**:6665–72.
78. Hsieh YL, et al. Salt stability—the effect of pHmax on salt to free base conversion. *Pharm Res* 2015;**32**(9):3110–18.
79. Anderson BD, Flora KP. . s.l. In: Wermuth CG, editor. *The practice of medicinal chemistry*. New York, NY: Academic Press; 1996. p. 732–54.
80. Guerrieri P, et al. Analysis of relationships between solid-state properties, counterion, and developability of pharmaceutical salts. s.l. *AAPS Pharm Sci Tech* 2010;**3**(11):1212–22..
81. Streng WH, et al. General treatment of pH-solubility profiles of weak acids and bases and effects of different acids on the solubility of a weak base. *J Pharm Sci* 1984;**73**(12):1679–84.
82. Mooney KG, et al. Dissolution kinetics of carboxylic acids I: effect of pH under unbuffered conditions. *J Pharm Sci* 1981;**70**:13–22.
83. Bogardus JB, Blackwood Jr. RK. Dissolution rates of doxycycline free base and hydrochloride salts. *J Pharm Sci* 1979;**68**(9):1183–4.
84. Li S, et al. Investigation of solubility and dissolution of a free base and two different salts. *Pharm Res* 2005;**22**(4):628–35.
85. *Preformulation study of a poorly water-soluble drug, α-pentyl-3-(2-quinolinylmethoxy) benzenemethanol: Selection of the base for dosage form design*. 5, 1986, J Pharm Sci, 75, pp. 492-496.
86. Chowhan ZT. pH-solubility profiles of organic carboxylic acids and their salts. *J Pharm Sci* 1978;**67**(9):1257–60.
87. Hawley M, Morozowich W. Modifying the diffusion layer of soluble salts of poorly soluble basic drugs to improve dissolution performance. *Mol Pharm* 2010;**7**(5):1441–9.
88. Badawy S, et al. Formulation of solid dosage forms to overcome gastric pH interaction of the factor Xa inhibitor, BMS-561389. *Pharm Res* 2006;**23**(5):989–96.
89. Ku SM. Use of the biopharmaceutical classification system in early drug development. *AAPS J* 2008;**10**(1):208–12.
90. Saal C, Becker A. Pharmaceutical salts: a summary on doses of salt formers from the orange book. *Eur J Pharm Sci* 2013;**49**:614–23.
91. Elder DP, et al. The utility of sulfonate salts in drug development. *J Pharm Sci* 2010;**99**(7):2948–61.
92. Presne C, et al. Lithium-induced nephropathy: rate of progression and prognostic factors. *Kidney Int* 2003;**64**(2):585–92.
93. Everett RM, et al. Nephrotoxicity of Pravadoline Maleate (WIN 48098-6) in dogs: evidence of maleic acid-induced acute tubular necrosis. *Fundamen Appl Toxicol* July 1993;**21**(1):59–65.
94. Olovson SG, et al. Oesophageal ulcerations and plasma levels of different alprenolol salts: potential implications for the clinic. *Acta Pharmacol Toxicol (Copenh)* 1986;**58**(1):55–60.
95. Thackaberry EA. Non-clinical toxicological considerations for pharmaceutical salt selection. *Exp Opin Drug Metab Toxicol* 2012;**8**(11):1419–33.
96. Badawy SI, Williams RC, Gilbert DL. Chemical stability of an ester prodrug of a glycoprotein IIb/IIIa receptor antagonist in solid dosage forms. *J Pharm Sci* 1999;**88**(4):428–33.
97. Zannou E, Joshi YM, Serajjudin ATM. Maleate salt of a basic drug by adjustment of microenvironmental pH in solid dosage forms. *Int J Pharm* 2007;**337**(1–2):210–18.
98. Guerrieri P, Taylor LS. Role of salt and excipient properties on disproportionation in the solid-state. *Pharm Res* 2009;**26**(8):2015–26.
99. John C, et al. Formulating weakly basic HCl salts: relative ability of common excipients to induce disproportionation and the unique deleterious effects of magnesium stearate. *Pharm Res* 2013;**30**(6):1628–41.
100. Stephenson GA, Aburub A, Woods TA. Physical stability of salts of weak bases in solid state. *J Pharm Sci* 2011;**100**(5):1607–17.
101. Hseih YL, Taylor LS. Salt stability—effect of particle size, relative humidity, temp and composition on salt to free base conversion. *Pharm Res* 2015;**32**:549–61.
102. Merritt JM, Viswanath SK, Stephenson GA. Implementing quality by design in pharmaceutical salt selection: a modeling approach to understanding disproportionation. *Pharm Res* 2013;**30**(1):203–17.
103. Serajuddin ATM, et al. Selection of solid dosage form composition through drug–excipient compatibility testing. *J Pharm Sci* 1999;**88**(7):696–704.
104. Schulthesiss N, Newman A. Pharmaceutical cocrystals and their physicochemical properties. *Crystal Growth Design* 2009;**9**(6):2950–67.
105. Good D, Rodriguez-Hornedo N. Solubility advantage of pharmaceutical cocrystals. *Crystal Growth Design* 2009;**9**(5):2252–64.
106. Nehm SJ, Rodrıguez-Spong B, Rodrıguez-Hornedo N. Phase solubility diagrams of cocrystals are explained by solubility product and solution complexation. *Crystal Growth Design* 2006;**6**(2):592–600.
107. Bak A, et al. The co-crystal approach to improve the exposure of a water-insoluble compound: AMG 517 sorbic acid co-crystal characterization and pharmacokinetics. *J Pharm Sci* 2008;**97**(9):3942–56.
108. Stahly GP. A Survey of cocrystals reported prior to 2000. *Crystal Growth Design* 2009;**9**:4212–29.
109. Thakuria R, et al. Pharmaceutical cocrystals and poorly soluble drugs. *Int J Pharm* 2013;**453**:101–25.
110. Good DJ, Rodriguez-Hornendo N. Cocrystal eutectic constants and prediction of solubility behavior. *Crystal Growth Design* 2010;**10**:1028–32.
111. Remenar JF, et al. Crystal engineering of novel cocrystals of a triazole drug with 1,4-dicarboxylic acids. *J Am Chem Soc* 2003;**125**(28):8456–7.
112. Childs SL, et al. Crystal engineering approach to forming cocrystals of amine hydrochlorides with organic acids. Molecular complexes of fluoxetine hydrochloride with benzoic, succinic, and fumaric acids. *J Am Chem Soc* 2004;**126**(41):13335–42.
113. Childs SL, Hardcastle KI. Cocrystals of piroxicam with carboxylic acids. *Crystal Growth Design* 2007;**7**(7):1291–304.
114. Lu E, Rodrıguez-Hornedo N, Suryanarayanan R. A rapid thermal method for cocrystal screening. *CrystEngComm* 2008;**10**:665–8.
115. Gagniere E, et al. Cocrystal formation in solution: Inducing phase transition by manipulating the amount of cocrystallizing agent. *J Cryst Growth* 2011;**316**:118–25.
116. Rodriguez-Spong B, et al. General principles of pharmaceutical solid polymorphism: a supramolecular perspective. *Adv Drug Deliv Rev* 2004;**56**:241–74.
117. Kulkarni GU, Kumardhas P, Rao CNR. Charge density study of the polymorphs of p-nitrophenol. *Chem Matter* 1998;**10**:3498–505.
118. Caira MR, et al. Structural characterization of two polymorphic forms of piroxicam pivalate. *J Pharm Sci* 1998;**87**:1608–14.
119. *International Conference on Harmonization Q6A Guideline. Specifications of New Drug Substances and Products: Chemical substances*. 1999 October.
120. Chen LR, et al. Solid-state behavior of cromolyn sodium hydrates. *J Pharm Sci* 1999;**88**:1191–200.

121. Burger A, Ramberger R. Polymorphism of pharmaceuticals and other molecular crystals. I. Theory of thermodynamic rules. *Microchim Acta* 1979;259–71.
122. Yu L. Inferring thermodynamic stability relationship of polymorphs from melting data. *J Pharm Sci* 1995;**84**:966–74.
123. Hu Y, et al. Estimation of the transition temperature for an enantiotropic polymorphic system from the transformation kinetics monitored using Raman spectroscopy. *J Pharm Biomed Anal* 2007;**45**(4):546–51.
124. Gu CH, Grant DJW. Estimating the relative stability of polymorphs and hydrates from heats of solution and solubility data. *J Pharm Sci* 2001;**90**(9):1277–87.
125. Liu C, et al. Facile method for the prediction of anhydrate/hydrate transformation point. *Ind Eng Chem Res* 2013;**52**:16506–12.
126. Khankari RK, Grant DJW. Pharmaceutical hydrates. *Thermochim Acta* 1995;**248**:61–79.
127. Tian F, et al. Factors affecting crystallization of hydrates. *J Pharm Pharmacol* 2010;**62**(11):1534–46.
128. Petrova RI, et al. Rapid conversion of API hydrates to anhydrous forms in aqueous media. *J Pharm Sci* 2009;**11**:4111–18.
129. Braun DE, et al. Navigating the waters of unconventional crystalline hydrates. *Mol Pharm* 2015;**12**(8):3069–88.
130. Yu L. Amorphous pharmaceutical solids preparation characterization and stabilization. *Adv Drug Deliv Rev* 2001; **48**:27–42.
131. Marsac PJ, Konno H, Taylor LS. A comparison of the physical stability of amorphous felodipine and nifedipine. *Pharm Res* 2006;**23**(10):2306–16.
132. Willart JF, Descamps M. Solid state amorphization of pharmaceuticals. *Mol Pharm* 2008;**5**(6):905–20.
133. Brittain H. *Polymorphism in pharmaceutical solids*, Vol. 95. New York, NY: Marcel Dekker; 1999.
134. Almarsson O, et al. High-throughput surveys of crystal form diversity of high polymorphic pharmaceutical compounds. *Crystal Growth Design* 2003;**3**(6):927–33.
135. Cui Y, Yao E. Evaluation of hydrate-screening methods. *J Pharm Sci* 2008;**97**(7):2730–44.
136. Zhu H, Yuen C, Grant DJ. Influence of water activity in organic solvent + water mixtures on the nature of the crystallizing drug phase. 1. Theophylline. *Int J Pharm* 1996;**135**:151–60.
137. Saunders M, et al. The potential of high speed DSC (Hyper-DSC) for the detection and quantification of small amounts of amorphous content in predominantly crystalline samples. *Int J Pharm* 2004;**274**(1–2):35–40.
138. Lefort R, et al. Solid state NMR and DSC methods for quantifying the amorphous content in solid dosage forms: an application to ball milling of trehalose. *Int J Pharm* 2004; **280**:209–19.
139. Saklatvala R, Royall PG, Craig DQ. The detection of amorphous material in a nominally crystalline drug using modulated temperature DSC—a case study. *Int J Pharm* 1999;**192**:55–62.
140. McGregor C, et al. The use of high-speed differential scanning calorimetry (Hyper-DSC™) to study the thermal properties of carbamazepine polymorphs. *Thermochimica Acta* 2004;**417**(2): 231–7.
141. McGregor C, Bines E. The use of high-speed differential scanning calorimetry (Hyper-DSC) in the study of pharmaceutical polymorphs. *Int J Pharm* 2008;**350**(1–2):48–52.
142. Gaisford S. Isothermal microcalorimetry for quantifying amorphous content in pharmaceuticals. *Adv Drug Deliv Rev* 2012;**64**:431–9.
143. Tishmack PA, Bugay DE, Byrn SR. Solid-state nuclear magnetic resonance spectroscopy—pharmaceutical applications. *J Pharm Sci* 2003;**92**(3):441–74.
144. Tong W-Q. Salt screening and selection: new challenges and considerations in the modern pharmaceutical research and development paradigm. In: Qiu Y, Chen Y, Zhang GZ, editors. *Developing solid oral dosage forms: pharmaceutical Theory and Practice*. 1st ed. Burlington, MA: Academic Press, Elsevier; 2009. p. 75–86.
145. Engel GL, et al. Salt form selection and characterization of LY333531 mesylate monohydrate. *Int J Pharm* 2000;**198**:239–47.
146. McNamara DP, et al. Use of a glutaric acid cocrystal to improve the oral bioavailability of low solubility API. *Pharm Res* 2006;**23**(8):1888–97.
147. Gao P, Shi Y. Characterization of supersaturatable formulations for improved absorption of poorly soluble drugs. *AAPS J* 2012;**14**(4):703–13.

CHAPTER

5

Drug Stability and Degradation Studies

D. Zhou[1], *W.R. Porter*[2] *and G.G.Z. Zhang*[1]

[1]AbbVie Inc., North Chicago, IL, United States [2]Abbott Laboratories, Abbott Park, IL, United States

5.1 INTRODUCTION

Degradation studies are an important aspect of preformulation evaluation of the stability of drug candidates. These experiments aim to aid further formulation development by evaluating the intrinsic stability properties of a drug candidate in a timely manner by deliberately applying stress to cause degradation. To fulfill this purpose, preformulation scientists induce degradation of test materials by raising temperature, by increasing humidity when relevant (eg, solid-state chemical and physical stability), by subjecting materials to sheer or compressive forces (eg, solid-state physical stability), by exposing the test materials to various pH conditions or intensive ultraviolet (UV)–visible light (eg, photostability) or by adding other reactants (eg, acids, bases, and peroxides). Forced degradation studies are sometimes described as stress testing, and they are also performed to support analytical method development.

Preformulation degradation studies and forced degradation studies used to support analytical method development share many common features. Analytical chemists employ forced degradation studies as part of their program to develop stability-indicating analytical methods. Early drug degradation studies also serve to identify and prepare degradants that are likely to appear in the final drug product and limit its shelf life. In addition, these studies aid in the elucidation of degradation pathways, as well as the qualification of materials used in toxicology experiments. Forced degradation studies conducted to support analytical method development often put more emphasis on identifying the degradants. Preformulation degradation studies, on the other hand, aim to predict the intrinsic stability of a drug candidate in order to anticipate problems that may arise down the development path. Cost savings can be realized if these activities can be coordinated.

Degradation studies can provide predictive information. Under certain circumstances, such as hydrolysis in solution, prediction of shelf life may be achieved with reasonable accuracy. However, in other cases, extrapolation of stability under stressed conditions (eg, oxidative testing) to normal storage conditions is challenging. Stress testing is, therefore, a more qualitative than quantitative predictive tool; nevertheless, rank-order comparisons are often not only possible, but also useful for designing subsequent formulations. Early signs of instability of a drug candidate help to identify potential development issues, to foster development of potential stabilization strategies, and to suggest ways to optimize manufacturing processes.

This chapter will discuss the basic treatments of drug degradation studies, including kinetics, pathways, important factors, and typical practices for assessing both the chemical and physical stability of pharmaceutical compounds.

5.2 CHEMICAL STABILITY

Chemical degradation probably represents the most important stability aspect of pharmaceuticals. Pharmaceutical scientists are responsible for examining the chemical stability of new drug candidates, for assessing the impact of stability issues on pharmaceutical development and processing, and for designing strategies to stabilize an unstable compound if necessary. They must understand the kinetics of chemical degradation, both in solution and in the solid state. They must also understand the commonly encountered

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00005-4

degradation pathways of active pharmaceutical ingredients (APIs), as well as practical approaches for performing degradation studies.

5.2.1 Solution kinetics

Chemical degradation reactions of pharmaceuticals follow the well-established treatments of chemical kinetics. A brief overview of chemical kinetics is presented next, followed by a discussion of catalysis and rate-pH profiles that are most relevant to the degradation of pharmaceuticals.

5.2.2 Rate equations

When a chemical reaction starts, the concentrations of reactants and products change with time until the reaction reaches completion or equilibrium. The concentrations of the reactants decrease, while those of the products increase over time. Therefore, the rate of a reaction can be represented either by the decreasing change in the concentration of a reactant or the increasing change in the concentration of a product with respect to time.

An arbitrary chemical reaction can be represented as

$$aA + bB \rightarrow cC + dD \tag{5.1}$$

Here, a, b, c, and d are the stoichiometric coefficients indicating the molar ratio of the reactants and products of the reaction. The rate of change of concentration of each species can differ, depending on the stoichiometric coefficients. Hence, a unified expression of the rate is preferred, which can be obtained via normalization:

$$\text{rate} = -\frac{1}{a}\frac{d[A]}{dt} = -\frac{1}{b}\frac{d[B]}{dt} = \frac{1}{c}\frac{d[C]}{dt} = \frac{1}{d}\frac{d[D]}{dt} \tag{5.2}$$

A negative sign is used for reactants so that the rate of a reaction is positive if it moves toward equilibrium or completion.

The rate of a reaction often depends on the concentrations of the reactants/products when other conditions are kept identical. Consider the hydrolytic reaction of ethyl acetate under alkaline conditions:

$$CH_3COOC_2H_5 + OH^- \rightarrow CH_3COO^- + C_2H_5OH \tag{5.3}$$

The rate of this reaction is proportional to the concentrations of each reactant species:

$$\begin{aligned}\text{rate} &= -\frac{d[CH_3COOC_2H_5]}{dt} = \frac{d[C_2H_5OH]}{dt} \\ &= k[CH_3COOC_2H_5][OH^-]\end{aligned} \tag{5.4}$$

Here, k, the proportional constant, is called the *specific rate constant*, or just the *rate constant*. This hydrolytic reaction is first order with respect to either ethyl acetate or hydroxide, and is an overall second-order reaction. The terms *first order* and *second order* will be defined later in this chapter.

In general, the rate of the arbitrary reaction (Eq. 5.1), may be written as

$$\text{rate} = k[A]^{\alpha}[B]^{\beta} \tag{5.5}$$

Here, α and β are the reaction order with respect to A and B, respectively. The order of the overall reaction is $n = \alpha + \beta$. This rate equation can be expanded to include more reactant/product species.

5.2.3 Elemental reactions and reaction mechanism

The way a reaction equation is written merely represents the overall outcome of the reaction and does not usually indicate how the actual reaction proceeds at a molecular level. For example, the reaction of hydrogen with bromine in the gas phase is usually written as

$$H_2 + Br_2 = 2HBr \tag{5.6}$$

As written, it appears as if one H_2 molecule collides with one Br_2 molecule and generates two HBr molecules in one step. However, the reaction actually proceeds with a free radical chain mechanism, which consists of the following steps:

$$Br_2 \rightarrow 2Br\bullet \tag{5.7}$$

$$H_2 + Br\bullet \rightarrow HBr + H\bullet \tag{5.8}$$

$$H\bullet + Br_2 \rightarrow HBr + Br\bullet \tag{5.9}$$

As shown previously, each step in a reaction scheme representing how the reactant molecules collide to form products is called an *elementary reaction*. The reaction equation representing the overall outcomes is called the *overall reaction*. The number of molecules reacting in an elementary reaction is called the *molecularity* of that elementary reaction, and this is a different concept from reaction order. An elementary reaction involving a single molecule is called *unimolecular*, one that involves two molecules is called *bimolecular*, and so on. Processes involving three (*termolecular*) or more molecules in one step are rare.

An overall reaction may contain a number of elementary reactions. A complete scheme of these elementary reactions represents the actual path of a reaction, and this is called the *mechanism* of the reaction. The beauty of the elementary reaction concept is that its rate follows the law of mass action. In simple words, the rate of an elementary reaction is directly proportional to the concentration of each reactant species involved in the elementary reaction, which is an outcome that depends upon the statistical aspects of

the reaction. Therefore, the rate equation of an elementary reaction can be readily derived.

Among the elementary reactions comprising an overall reaction, one of these steps usually proceeds with the slowest rate. This slowest step is called the *rate-determining* step of a reaction and determines the rate of the overall reaction. Once the reaction mechanism is known and the rate-determining step is identified, the expression of the overall rate of the reaction may also be readily derived.

5.2.4 Typical simple order kinetics

Simple order reactions described in the following sections refer to reactions that are zero-, first-, and second-order. These reaction orders are commonly encountered in studies of the stability of pharmaceuticals. Their mathematical treatments are also relatively simple.

5.2.4.1 Zero-order reactions

In zero-order reactions, the rate of the reaction does not depend on the concentration of the reactant; thus, the rate is a constant:

$$\text{rate} = -\frac{d[\text{A}]}{dt} = k[\text{A}]^0 = k \tag{5.10}$$

Here, A is the reactant and k is the zero-order rate constant. In this case, the decrease in concentration of A is linear with time:

$$[\text{A}]_t = [\text{A}]_0 - kt \tag{5.11}$$

Here, $[\text{A}]_t$ is the concentration of A at time t, while $[\text{A}]_0$ is that at time zero, or the initial concentration.

5.2.4.2 First-order reactions

First-order reactions appear to be the most commonly encountered in pharmaceutical stability studies. The rate of a first-order reaction is proportional to the concentration of the reactant:

$$\text{rate} = -\frac{d[\text{A}]}{dt} = k[\text{A}] \tag{5.12}$$

The concentration-time profile of the reactant for a first-order reaction follows an exponential decay to a limiting value, while that of the product follows an exponential increase to a different limiting value:

$$\text{A} \rightarrow C \tag{5.13}$$

$$[\text{A}]_t = [\text{A}]_0 \exp(-kt) \tag{5.14}$$

$$[\text{C}]_t = [\text{A}]_0[1 - \exp(-kt)] \tag{5.15}$$

The half-life, $t_{1/2}$, of the reaction is the time required for the reactant concentration to decrease to 50% of its original value; similarly, the times for the reactant concentration to decrease to 95% and 90% of its original values are designated as t_{95}, and t_{90}, respectively. These quantities can be obtained readily for a first-order reaction if the rate constant is known:

$$t_{1/2} = \frac{\ln 2}{k}; \quad t_{95} = \frac{\ln 0.95}{k}; \quad t_{90} = \frac{\ln 0.9}{k} \tag{5.16}$$

A characteristic feature of first-order reactions is that the time required to lose the first 50% of the material ($t_{1/2}$) is the same as the time required to drop from 50% remaining to 25% remaining, from 25% remaining to 12.5% remaining, and so on.

5.2.4.3 Second-order reactions

Many apparently first-order reactions observed for pharmaceuticals are actually second order. Usually, two reactant molecules must collide in order to react. However, in practice, one reactant (eg, water, hydrogen ion, hydroxyl ion, buffer species, etc.) may be in great excess so that its change in concentration is negligible, and an apparent first-order reaction is therefore observed.

For a second-order reaction where two reactants are involved,

$$\text{A} + \text{B} \rightarrow \text{C} \tag{5.17}$$

The rate equation can be written as

$$\text{rate} = -\frac{d[\text{A}]}{dt} = -\frac{d[\text{B}]}{dt} = k[\text{A}][\text{B}] \tag{5.18}$$

The rate is first-order with respect to each reactant, but the overall reaction is second order. The concentration-time profile of a second-order reaction can be represented as

$$\frac{1}{[\text{A}]_0 - [\text{B}]_0}\left(\ln\frac{[\text{A}]_t}{[\text{B}]_t} - \ln\frac{[\text{A}]_0}{[\text{B}]_0}\right) = kt \tag{5.19}$$

When the initial concentrations of A and B are identical, the concentration-time profile can be simplified as

$$\frac{1}{[\text{A}]_t} - \frac{1}{[\text{A}]_0} = kt \tag{5.20}$$

The $t_{1/2}$, t_{95}, and t_{90} values for a second-order reaction all depend upon the initial concentration of each species.

Fig. 5.1 plots the reactant concentration-time profiles for theoretical zero-, first-, and second-order kinetics. Table 5.1 summarizes the rate equations, the formula for calculating reactant concentration-time profiles, and half-lives for these simple order kinetics. The rate constants used to generate Fig. 5.1 were assumed to be numerically identical in all cases. Identical initial reactant concentrations were assumed for the second-order reaction in both Fig. 5.1 and Table 5.1.

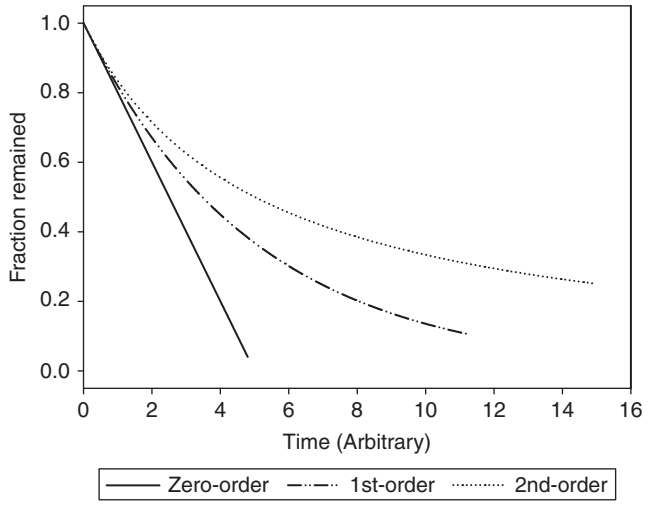


FIGURE 5.1 Reactant concentration-time profiles for theoretical zero-, first-, and second-order reactions.

TABLE 5.1 Rate Equations, Reactant Concentration-Time Profiles, and Half-Lives for Zero-, First-, and Second-Order Reactions

Reaction order	Rate equation	Concentration-time profile	Half-life
Zero	$-\frac{d[A]}{dt} = k$	$[A]_t = [A]_0 - kt$	$t_{1/2} = \frac{[A]_0}{2k}$
First	$-\frac{d[A]}{dt} = k[A]$	$[A]_t = [A]_0 \exp(-kt)$	$t_{1/2} = \frac{\ln 2}{k}$
Second	$-\frac{d[A]}{dt} = k[A]^2$	$\frac{1}{[A]_t} - \frac{1}{[A]_0} = kt$	$t_{1/2} = \frac{1}{k[A]_0}$

5.2.4.4 *Apparent pseudokinetic orders*

Zero-, first-, and second-order kinetics have been briefly discussed already in this chapter; these models are likely to be the most frequently encountered in drug stability studies. In practice, only the concentration of the API is usually monitored over time, while other reactants such as hydrogen ion, hydroxyl ion, and other buffer components are in great excess and are usually kept constant by methods such as pH-stating. Hence, a second-order overall reaction could appear to be a first-order reaction (with respect to drug concentration). The reaction is now said to be apparently first-order or pseudo-first-order. This is the case for most pharmaceutical hydrolytic reactions that are catalyzed by various species of buffers.

Apparent zero-order kinetics arise in situations when a source of the reactant exists that maintains the concentration of the reactant during the reaction. For example, the concentration of a drug in the solution phase of a suspended formulation is kept constant (at its solubility limit) if the dissolution rate of the drug from the solid phase is faster than the rate of degradation. Supposing that only the drug molecules in solution degrade significantly, the overall degradation kinetics would be apparently zero-order or pseudo-zero-order, although the actual kinetics may be first-order or apparently first-order in solution. Apparent zero-order kinetics may also arise in other situations, such as degradation in a surfactant solution where the overall concentration of the surfactant is much higher than the critical micelle concentration, and degradation in the micellar phase is negligible. Finally, apparent zero-order kinetics occur when the extent of degradation is so small that the dependence of the reaction rate on the concentration of a reactant cannot be demonstrated because of experimental errors in measurement. Fig. 5.1 shows that, assuming a reaction to proceed by a zero-order mechanism always predicts the greatest degree of decomposition as a function of time, other factors being equal. Thus, when the order of a reaction is in doubt, the conservative choice is to assume a zero-order kinetic model for the purpose of predicting the maximum amount of degradation that could be expected at some future time.

A distinction between true and apparent kinetic orders can be made by changing concentrations of other species (eg, buffer concentration, pH, or even different buffer components) or by changing the system so that the solubility of a drug can be altered without changing the reaction mechanism.

5.2.5 Complex reactions

Many reactions involve more than a single step and are known as *complex reactions*. Depending on the reaction schemes and the magnitude of respective rate constants, the overall kinetics may be approximated by zero-, first-, or second-order rate equations. However, more often than not, the kinetic expressions are more complicated. Some commonly encountered complex reaction schemes are described next.

5.2.5.1 *Reversible reactions*

Essentially, all reactions are reversible to some extent. However, the equilibrium constant may be so huge that the overall reaction can be treated as virtually one-directional. For the following simplest reversible reaction:

$$A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} B \tag{5.21}$$

Here, k_1 and k_{-1} are the first-order rate constants for the forwarding and reversing reactions, respectively. The rate equation can be written as

$$\text{rate} = -\frac{d[A]}{dt} = \frac{d[B]}{dt} = k_1[A] - k_{-1}[B] \tag{5.22}$$

At equilibrium, the concentrations of species A and species B do not change any further because the

rate of the forward reaction equals that of the reverse reaction:

$$\text{rate} = -\frac{d[\text{A}]}{dt} = \frac{d[\text{B}]}{dt} = k_1[\text{A}]_{\text{eq}} - k_{-1}[\text{B}]_{\text{eq}} = 0 \qquad (5.23)$$

The subscript "eq" designates quantities at equilibrium. Therefore:

$$K = \frac{[\text{B}]_{\text{eq}}}{[\text{A}]_{\text{eq}}} = \frac{k_1}{k_{-1}} \qquad (5.24)$$

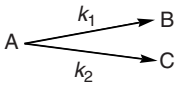
K is the equilibrium constant and is equal to the ratio between the rate constant of the forward reaction to that of the reverse reaction. Thus, Eq. (5.24) provides a connection between the thermodynamic chemical equilibrium and the kinetics of a reversible reaction. The concentration-time profile of a reversible first-order reaction can then be obtained:

$$\ln\left(\frac{[\text{A}]_0 - [\text{A}]_{\text{eq}}}{[\text{A}] - [\text{A}]_{\text{eq}}}\right) = (k_1 + k_{-1})t \qquad (5.25)$$

Fig. 5.2 shows the time-concentration profiles for a theoretical reversible reaction, assuming $K = 3$, $k_1 = 0.3$, and $k_{-1} = 0.1$.

5.2.5.2 Parallel reactions

Sometimes more than one reaction pathway exists for degradation of a drug candidate. These pathways might lead to identical degradants or, more commonly, different degradants. Take for an example the following simplest parallel reaction:



(5.26)

Here k_1 and k_2 are the first-order rate constants for reaction A→B and A→C, respectively. The corresponding rate equation is:

$$\text{rate} = -\frac{d[\text{A}]}{dt} = k_1[\text{A}] + k_2[\text{A}] = (k_1 + k_2)[\text{A}] = k_{\text{obs}}[\text{A}] \qquad (5.27)$$

Here, $k_{\text{obs}} = k_1 + k_2$ is the observed apparent first-order rate constant. Mathematically, the concentration-time profiles can be obtained as follows:

$$[\text{A}]_t = [\text{A}]_0 \exp(-k_{\text{obs}}t) \qquad (5.28)$$

$$[\text{B}]_t = [\text{B}]_0 + \frac{k_1}{k_{\text{obs}}}[\text{A}]_0\left[1 - \exp(-k_{\text{obs}}t)\right] \qquad (5.29)$$

$$[\text{C}]_t = [\text{C}]_0 + \frac{k_2}{k_{\text{obs}}}[\text{A}]_0\left[1 - \exp(-k_{\text{obs}}t)\right] \qquad (5.30)$$

Fig. 5.3 shows the concentration-time profile for a parallel reaction with $k_1 = 0.2$ and $k_2 = 0.1$.

5.2.5.3 Consecutive reactions

Consecutive reactions involve an intermediate formed from the initial reactant that in turn is converted to the final product. For a simple first-order consecutive reaction from reactant A to product C via intermediate B:

$$A \xrightarrow{k_1} B \xrightarrow{k_2} C \qquad (5.31)$$

The rate equations may be obtained as

$$-\frac{d[\text{A}]}{dt} = k_1[\text{A}] \qquad (5.32)$$

$$\frac{d[\text{B}]}{dt} = k_1[\text{A}] - k_2[\text{B}] \qquad (5.33)$$

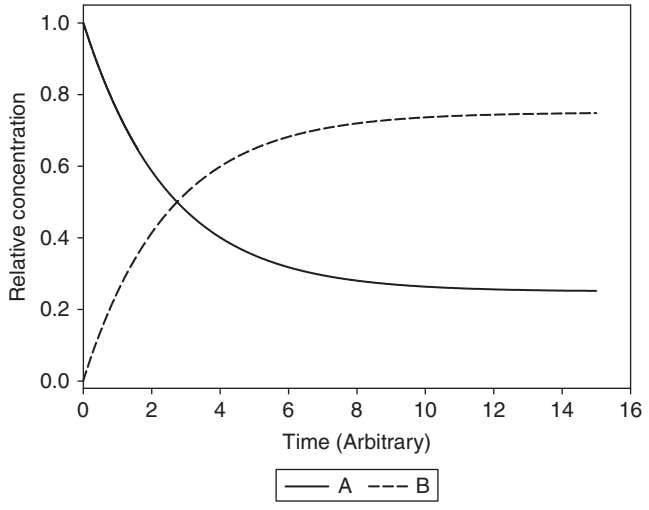


FIGURE 5.2 Concentration-time profiles for a theoretical reversible reaction.

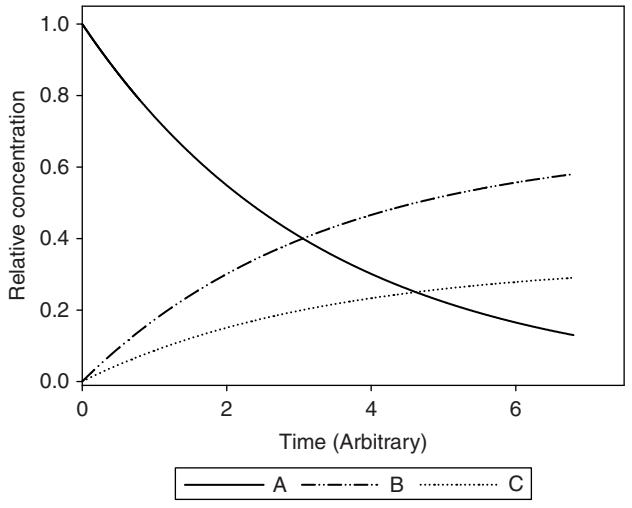


FIGURE 5.3 Concentration-time profiles for a theoretical parallel reaction.

$$\frac{d[\mathrm{C}]}{dt} = k_2[\mathrm{B}] \tag{5.34}$$

The concentration-time profiles can be obtained by simultaneously solving the differential equations (Eqs. 5.32–5.34), getting the following results:

$$[\mathrm{A}] = [\mathrm{A}]_0 \exp(-k_1 t) \tag{5.35}$$

$$[\mathrm{B}] = \frac{k_1}{k_2 - k_1}[\mathrm{A}]_0\{\exp(-k_1 t) - \exp(-k_2 t)\} \tag{5.36}$$

$$[\mathrm{C}] = [\mathrm{A}]_0 + \frac{1}{k_1 - k_2}[\mathrm{A}]_0\{k_2 \exp(-k_1 t) - k_1 \exp(-k_2 t)\} \tag{5.37}$$

The concentration-time profiles for a theoretical consecutive reaction are shown in Fig. 5.4 for a theoretical consecutive reaction A→B→C with $k_1 = 0.2$, and $k_2 = 0.1$. Note that the time profile of the intermediate B shows a maximum.

5.2.6 Arrhenius equation, collision theory, and transition state theory

5.2.6.1 *Arrhenius equation*

The rate of a reaction is usually affected by concentration, solvents, and catalysis and, very importantly, by temperature. The Q10 rule, attributed to pioneering physical chemist Jacobus H. van't Hoff, states that the velocity of a chemical reaction increases approximately twofold to threefold for each 10°C rise in temperature. His student, Svante Arrhenius, expressed this more formally as[1]

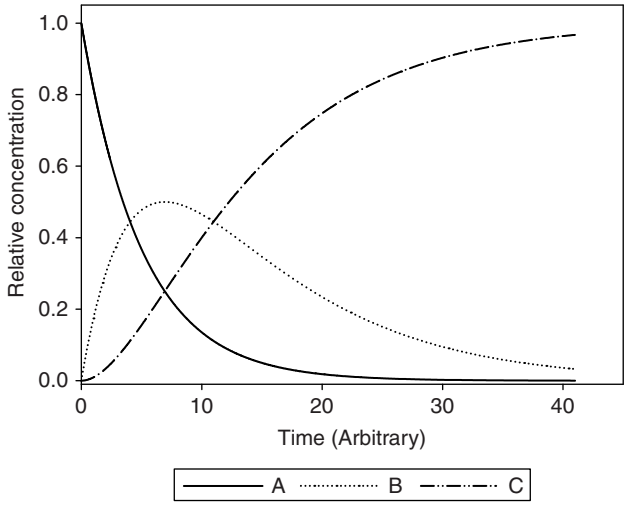


FIGURE 5.4 Concentration-time profiles for a theoretical consecutive reaction.

$$\frac{d\ln(k)}{dT} = \frac{E_a}{RT^2} \tag{5.38}$$

This can be integrated to give the exponential form:

$$\ln(k) = \ln(A) - \frac{E_a}{RT} \quad \text{or} \quad k = A\exp\left(-\frac{E_a}{RT}\right) \tag{5.39}$$

Here, k is the rate constant, A is a constant known as the *Arrhenius frequency factor* (where the limiting rate as the absolute temperature is increased to infinity), E_a is another constant called the *activation energy*, R is the *universal gas constant*, and T is the *absolute temperature*. The frequency factor determines the collision frequency between reacting molecules, while the activation energy is the energy barrier that the molecules must overcome in order to react and form products. Experimentally, the frequency factor and the activation energy can be obtained from the intercept and the slope, respectively, of the Arrhenius plot of $\ln(k)$ versus $1/T$.

Connors et al. pointed out that most drug degradations have activation energies of 12–24 kcal/mol[2]; however, higher activation energies have also been observed.[3] The Q10 rule assumes an activation energy of 13–20 kcal/mol; therefore, it provides a conservative estimate of the effect of temperature on drug degradation rates.

5.2.6.2 *Classic collision theory of reaction rates*

Reactant molecules must collide in order to react. However, not every collision leads to a successful reaction. The classical collision theory of reaction rates postulates that the reaction takes place only if the involved molecules exceed a certain minimum energy (the activation energy). Based on the abovementioned assumptions, an expression of the rate similar to the form in the Arrhenius equation can be derived.

The fraction of molecules with kinetic energies exceeding E_a can be derived using the Boltzmann distribution:

$$f(E_k > E_a) = \exp\left(-\frac{E_a}{RT}\right) \tag{5.40}$$

Here, f is the fraction reacting (or probability of reaction), and E_k is the kinetic energy. Therefore, the rate can be expressed as

$$\text{rate} = ZPC_{E_k > E_a} = ZPC_T \exp\left(-\frac{E_a}{RT}\right) \tag{5.41}$$

Here, $C_{E_k > E_a}$ is the concentration of the reactant molecules with kinetic energy exceeding the activation energy, C_T is the total reactant concentration, Z is the collision frequency, and P is the steric factor ($0 < P < 1$), because not every collision leads to a

reaction. By comparison with the general rate law, one can immediately see that the expression of the rate constant, k, is similar to the Arrhenius equation:

$$k = ZP \exp\left(-\frac{E_a}{RT}\right) \tag{5.42}$$

Here, ZP is identified with the Arrhenius frequency factor A, while E_a is the activation energy.

5.2.6.3 *Transition state theory*

The classic collision theory assumes that the reaction is possible when reactant molecules with sufficient energies collide. It does not answer the question of how they react. The transition state theory, developed around 1935 by Eyring and Polanyi et al. (see the review by Eyring[4]) postulates that in addition to collision, a transition state (or activated complex) must exist in which the reacting molecules have an appropriate configurational geometry for the reaction to occur; furthermore, the transition state is in equilibrium with the reactants. In order for the activated complex to form, a certain energy barrier (the activation energy) needs to be overcome. Once the activated complex is formed, it can decompose to products. In principle, the transition theory provides a means to calculate the rate of a reaction once certain properties of the activated complex, such as vibrational frequency, molecular mass, and interatomic distances, are known. This theory is also referred to as the *absolute rate theory*.

A simplified scheme can be represented as follows:

$$\mathrm{A} + \mathrm{B} \overset{K^{\ddagger}}{\rightleftharpoons} [A\ldots B]^{\ddagger} \rightarrow \mathrm{C} \tag{5.43}$$

The double dagger designates the transition state.

Transition state theory assumes that the decomposition of the activated complex is the rate-determining step of the reaction. Among all the molecular motions of the activated complex, there is one mode of vibration (frequency, designated as ν) that will lead to the formation of the product. Combined with the equilibrium assumption, the rate of the reaction can be written as

$$\text{rate} = \frac{d[\mathrm{C}]}{dt} = \nu[\mathrm{A}\cdots\mathrm{B}]^{\ddagger} = \nu K^{\ddagger}[\mathrm{A}][\mathrm{B}] \tag{5.44}$$

Thus, the rate constant of the reaction, k, has the following form:

$$k = \nu K^{\ddagger} \tag{5.45}$$

Once the vibration frequency ν and the equilibrium constant are known, the rate constant can be calculated. For example, the equilibrium constant can be obtained from statistical thermodynamic treatments of chemical equilibrium, while the frequency of the decomposition vibration can, as shown by Eyring, be approximated by

$$\nu = \frac{k_B T}{h} \tag{5.46}$$

Here, k_B is Boltzmann's constant, h is Plank's constant, and T is the absolute temperature. The final rate constant can be expressed as

$$\begin{aligned} k &= \frac{k_B T}{h} K^{\ddagger} = \frac{k_B T}{h} \exp\left(-\frac{\Delta G^{\ddagger}}{RT}\right) \\ &= \frac{k_B T}{h} \exp\left(\frac{\Delta S^{\ddagger}}{R}\right) \exp\left(-\frac{\Delta H^{\ddagger}}{RT}\right) \end{aligned} \tag{5.47}$$

$\Delta G^{\ddagger}$, $\Delta S^{\ddagger}$, and $\Delta H^{\ddagger}$ are the free energy, entropy, and enthalpy of activation, respectively. In comparison with the Arrhenius equation, it is immediately obvious that

$$\begin{aligned} E_a &= RT^2 \frac{d \ln k}{dT} = RT + \Delta U^{\ddagger} \\ &= RT + \Delta H^{\ddagger} - \Delta(pV^{\ddagger}) \cong RT + \Delta H^{\ddagger} \end{aligned} \tag{5.48}$$

$$A = \frac{k_B T}{h} \exp\left(\frac{\Delta S^{\ddagger}}{R}\right) \tag{5.49}$$

$\Delta(pV^{\ddagger})$ is the change of pV when the activated complex is formed, and this value is negligible for most condensed phase reactions.

The transition state theory provides a better picture of chemical reactivity, and it is still used today. It accounts for the influences on reaction rate by various factors, such as solvent, ionic strength, and dielectric constant. In general, enhancement of the reaction rate is expected if a factor stabilizes the transition state. However, sometimes changing one factor might change the effects of other factors. Therefore, caution must be taken when such exercises are performed.

A schematic representation of the transition state theory is shown in Fig. 5.5a. Fig. 5.5b describes the situation where an intermediate exists. The intermediate is located at a local minimum on the potential energy surface, and in this case, there are two separate activation energies.

5.2.7 Catalysts and catalysis

A *catalyst* is a substance that influences the rate of a reaction without itself being consumed. Positive catalysis increases the reaction rate, while negative catalysis reduces the reaction rate. However, positive catalysis is what is normally considered.

A catalyst changes the rate of a reaction by altering the free energy of activation (or activation energy)

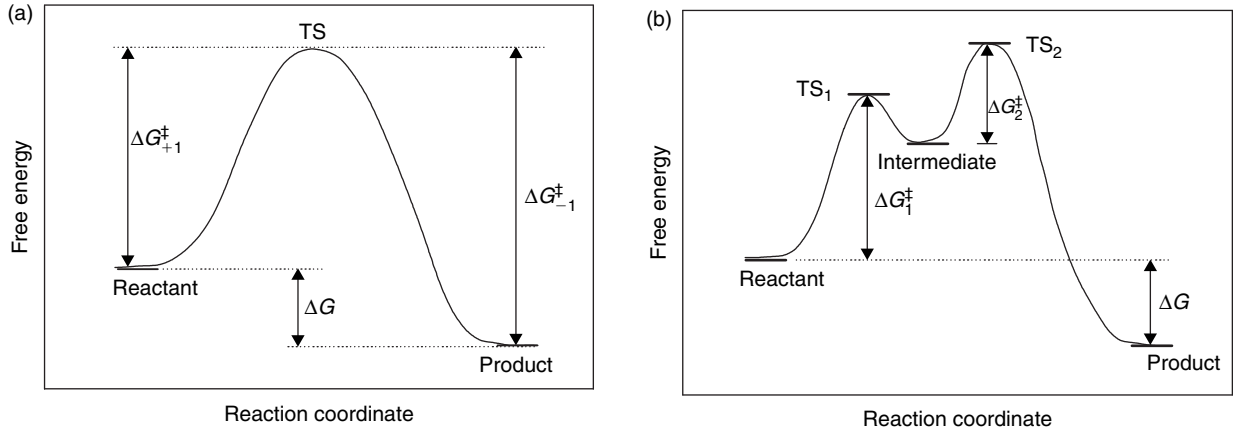


FIGURE 5.5 Schematic potential energy surface along the reaction coordinates; (a) single-step reaction; (b) reaction with an intermediate.

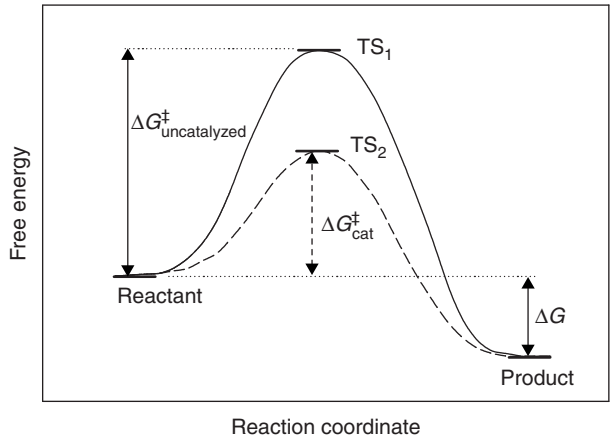


FIGURE 5.6 Schematic potential energy surfaces along a reaction coordinate with a positive catalyst.

without altering the thermodynamic aspects of the reaction. In other words, the equilibrium of a reaction is not changed by the presence of a catalyst. A positive catalyst reduces the free energy of activation, while a negative catalyst increases it. Fig. 5.6 describes the reaction coordinate with a positive catalyst.

The activation energy decreases in the presence of the catalyst; however, the thermodynamics, as indicated by the ΔG of the reaction, are not altered.

5.2.7.1 *Specific acid-base catalysis*

Hydrogen ions and hydroxyl ions are often involved directly in the degradation of pharmaceuticals. In many cases, the concentration of hydrogen ions or hydroxyl ions appears in the rate equation; the reaction is then said to be subject to specific acid-base catalysis.

For example, an ester drug is hydrolyzed in an acidic buffer; the second-order rate equation is

$$\text{rate} = -\frac{d[\text{D}]}{dt} = k\left[\text{H}^+\right][\text{D}] \tag{5.50}$$

Drug degradation studies are often carried out in buffered solutions where concentrations of hydrogen or hydroxyl ions are maintained as constant. In addition, the drug itself is usually the only species whose concentration is monitored over time. As a result, the degradation kinetics is pseudo-first-order when the reaction is catalyzed by hydrogen or hydroxyl ions, with the apparent rate constant, k_{obs}:

$$k_{obs} = k[\text{H}^+] \quad \text{or} \quad \log k_{obs} = \log k - \text{pH} \tag{5.51}$$

For a reaction subject to specific acid catalysis, a plot of the logarithm of the apparent first-order rate constant with respect to pH gives a straight line of slope -1.

Similarly, for specific base catalysis, a plot of the logarithm of the apparent first-order rate constant versus pH gives a line of slope of $+1$:

$$k_{obs} = k[\text{OH}^-] \quad \text{or} \quad \log k_{obs} = \log k + \text{pH} \tag{5.52}$$

Hydrolysis of zileuton is subject to specific acid catalysis, but not base catalysis.[5,6] Hydrolysis of aspirin shows both specific acid- and specific base-catalysis.[7]

5.2.7.2 *General acid-base catalysis*

General acid-base catalysis occurs when the buffering components catalyze a reaction. Either the acidic or basic components of the buffer, or both, can catalyze a reaction.

General acid-base catalysis often causes deviation of the rate-pH profile from the expected behavior. The exact magnitudes of general acid-base catalysis can often be obtained by varying buffer concentration while keeping the remaining conditions (ionic strength,

pH, etc.) constant. A plot of the apparent first-order rate constant versus buffer concentration should be a straight line whose slope corresponds to the catalytic coefficient of the buffer at the studied pH. The intercept of this plot corresponds to the specific rate constant in the absence of general acid-base catalysis.

No distinction between the species responsible for general catalysis can be made by this method. However, it could be achieved if the reaction is studied at different pH values with the same buffer species.

In general, the apparent first-order rate constant can be written as follows, taking into account both specific and general acid-base catalysis:

$$\begin{aligned} k_{obs} &= k_0 + k_H[H^+] + k_{OH}[OH^-] + k_1[\text{buffer species 1}] \\ &\quad + k_2[\text{buffer species 2}] + \ldots \\ &= k_0 + \sum_i k_i C_i \end{aligned} \tag{5.53}$$

Here, k_0 is the intrinsic first-order rate constant in the absence of any catalysis, C_i is the concentration of species i, and k_i is the catalytic coefficient of species i.

The degradation of a number of pharmaceuticals has been shown to be subject to general acid-base catalysis. Degradation of thiamine was found to be catalyzed by acetate ion, but not acetic acid, and is therefore an example of general base catalysis.[8] Citric acid, citrate ions, and phosphate ions all catalyze the decomposition of ampicillin, an example illustrating both general acid and general base catalysis.[9]

5.2.8 pH-rate profiles

The pH-rate profile, sometimes called the *pH-stability profile* or *rate-pH profile*, is the pH dependence of the specific rate constant of degradation of a compound; it is conveniently represented by a log(k) versus pH plot. The pH-rate profiles provide insights into the catalytic nature of a reaction, and they can help in developing more stable solution formulations and lyophilized products. In addition, pH-rate profiles aid in the development of more stable conventional solid oral dosage forms (eg, tablets), although solution stability information is not required to be as detailed in this case.

Many drug degradation reactions follow apparent first-order kinetics and are subject to specific and general acid-base catalysis. The apparent first-order rate constants are usually plotted in a pH-rate profile. One should correct for general acid-base catalysis by buffer components by extrapolation to zero buffer concentration if the catalysis effect is significant.

Analysis of a pH-rate profile can be started by assuming all possible pathways and writing down the corresponding rate equations (eg, Eq. 5.53). The presence or absence of a certain mechanism can then be verified by analyzing the kinetic data. A brief discussion of the more basic pH-rate profiles follows.

5.2.8.1 *V-shaped, U-shaped, and other truncated pH-rate profiles*

Specific acid and base catalysis is common in the degradation of carboxylic acid derivatives, such as esters, amides, substituted ureas, etc. In the absence of other more complicated mechanisms, the pseudo-first-order rate constant can be written as

$$k_{obs} = k_H[H^+] + k_0 + k_{OH}[OH^-] \tag{5.54}$$

Here, k_0 is the intrinsic apparent first-order rate constant, and k_H and k_{OH} are the catalytic coefficients for the hydrogen and hydroxyl ions, respectively. The resulting rate-pH profile plot consists of a straight line with slope of -1 in the acidic region, and another straight line with slope of 1 in the basic region (Fig. 5.7a). Atropine[10,11] and diazepam[12] are examples of compounds with V-shaped pH-rate stability profiles. When k_0 is sufficiently large, the V-shaped pH-rate profile may show a flat region and the resulting plot is U-shaped (see Fig. 5.7b; hydrolysis of the β-lactam ring in cephalothin[13]).

Degradation of some drug substances may show only specific acid- or base-catalysis, but not both (eg, hydrolysis of echothiophate[14]). The plot (Fig. 5.7c or 5.7d) then either has a slope of -1 in the acidic region joined to a flat line extending into the basic region (specific acid catalysis) or a flat line in the acidic region joined to a line with slope 1 in the basic region (specific base catalysis). In some cases, only specific acid-catalysis or specific base-catalysis is observed; therefore, the rate-pH profile consists of only a straight line (Fig. 5.7e and 5.7f; eg, lansoprazole[15]). Fig. 5.7 exhibits these rate-pH profiles where only a single species is reactive with either specific acid- or base-catalysis, or both.

5.2.8.2 *Sigmoidal pH-rate profiles*

Sigmoidal rate-pH profiles usually arise from the dissociation of the drug molecules. Many drug candidates are weak bases or weak acids. In the vicinity of $pH = pK_a$, the species distributions of a weak base or weak acid are sigmoidal when plotted as a function of pH. Therefore, if both the acidic and basic species of the compound can undergo degradation with differing rate constants, the rate-pH profile is expected to be sigmoidal. For example, for the decomposition of weak acid HA:

$$HA \xrightarrow{k_{HA}} \text{product} \tag{5.55}$$

$$A^- \xrightarrow{k_{A^-}} \text{product} \tag{5.56}$$

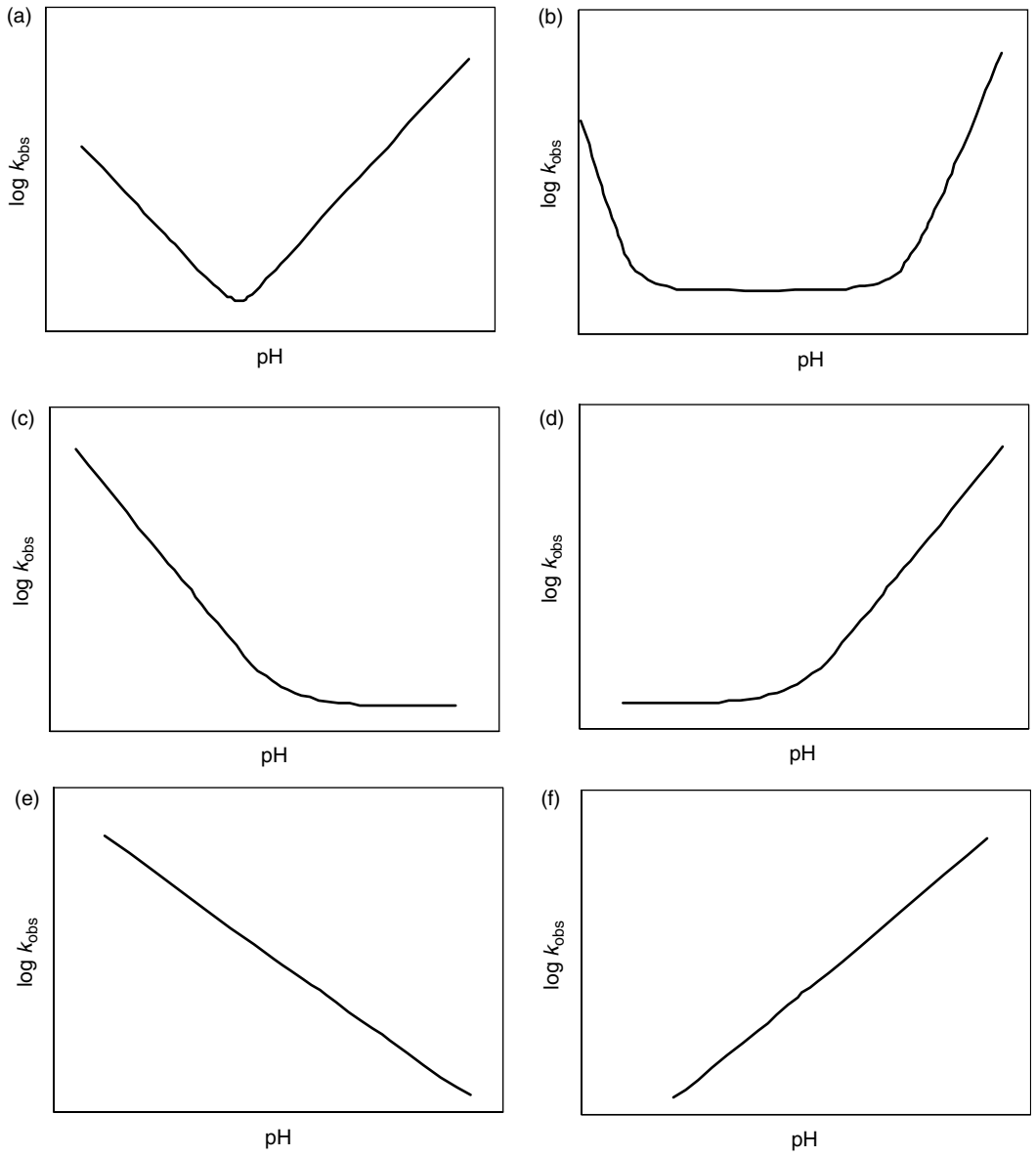


FIGURE 5.7 Schematic rate-pH profiles for reactions involving only a single reactive species with specific acid-, base-catalysis, or both.

When the drug concentration is measured, a distinction between the ionized and unionized species is usually not made. The apparent rate of the reaction is

$$\begin{aligned} \text{rate} &= k_{HA}[HA] + k_{A^-}[A^-] \\ &= \frac{k_{HA}[H^+] + k_{A^-}K_a}{K_a + [H^+]}\{HA\} \end{aligned} \tag{5.57}$$

Here, K_a is the dissociation constant of HA, while {HA} is the total concentration of HA. Therefore, a plot of the apparent rate constant is sigmoidal with respect to pH, so long as the rate constants are not identical. The rate constant of each species can be estimated from the limits of the apparent rate constant at low and high pH, and that pK_a = pH at the inflection point of the sigmoidal pH-rate profile plot. If the change in rate is due to ionization at a single site, the sigmoidal curve will encompass slightly more than ± 1 pH units of the expected pK_a.

In this instance, 5-flurouracil is a weak diprotic acid and has 2 ionization constants with $pK_1 = 7.3$ and $pK_2 = 11.3$. Among the three species, A^{2-} has the highest reactivity, while the neutral molecules H_2A and HA^- possess only slightly different rate constants. Therefore, the sigmoidal pH-rate profile[16] is rather obvious in the pH range 10.5–12.5 (Fig. 5.8a). While the change in rate from pH 6 to pH 8 is only minor, the sigmoid shape can still be seen around the first pK_a.[16] Cycloserine is a zwitterion, with $pK_1 = 4.5$ (acid) and $pK_2 = 7.4$ (base). The three species have different

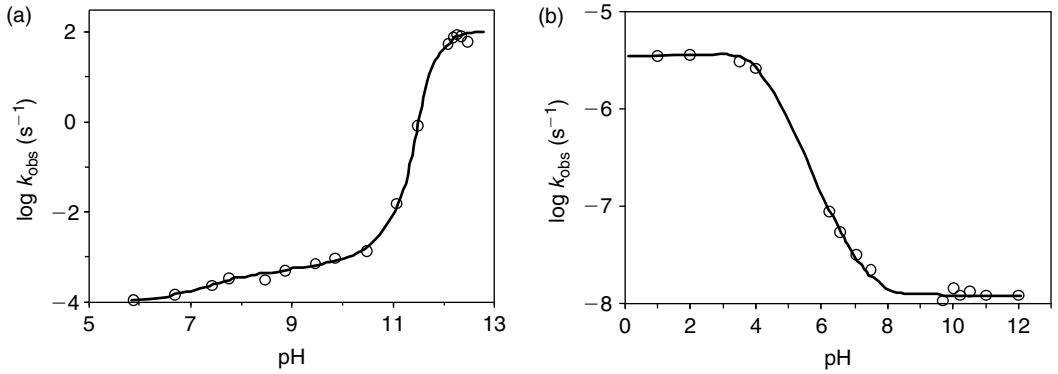


FIGURE 5.8 Sigmoidal pH-rate profiles; (a) 5-fluorouracil; (b) cycloserine at 25°C. *Redrawn from Garrett ER, Nestler HJ, Somodi A. Kinetics and mechanisms of hydrolysis of 5-halouracils.* J Org Chem *1968;**33**:3460–8. Kondrat'eva AP, Bruns BP, Libinson GS. Stability of D-cyclo-serine in aqueous solutions at low concentrations.* Khimiko-Farmatsevticheskii Zhurnal *1971;**5**:38–41, respectively.*

reactivity and follow the following order: $AH^{2+} > AH^{\pm} > A^{-}$. Therefore, a broad sigmoidal pH-rate profile encompassing the pH range 3–9 (Fig. 5.8b) was observed.[17] Degradation of barbital[18] and zileuton[6] also shows similar behavior.

5.2.8.3 Bell-shaped pH-rate profiles

Bell-shaped pH-rate profiles show a minima or maxima. Different scenarios can lead to this kind of pH-rate profile. The most intuitive scenario arises from the existence of two ionizable functional groups in the molecule. For example, for a diprotic acid, H_2A, three species are in solution: H_2A, HA^{-}, and A^{2-}, where the concentration-pH profile of species HA^{-} is bell-shaped. Depending on whether the monoprotic species, HA^{-}, is the most or least reactive, and the corresponding pH-rate profile could show either a maxima or a minima. Based on a similar argument, one can also show that the reaction between an acid and a base can lead to a bell-shaped pH-rate profile. In this case, the two ionizations are on different reactant molecules.

The third scenario occurs when ionization is combined with a change in the rate-determining step. For example, consider a consecutive reaction: A→B→C, where A is a monoprotic acid/base. The two species of reactant A may have very different reactivities with the rate constant of step B→C falling somewhere in between. Therefore, in one pH region (below or above its pK_a), the step A→B is the slowest, whereas B→C becomes the rate-determining step over another pH range. A bell-shaped pH-rate profile then results, with one side of the bell corresponding to the ionization while the other corresponds to the switch of the rate-limiting step. Hydrolysis of hydrochlorothiazide[19,20] was proposed to be such a case (Fig. 5.9). Its profile from pH 7 to 12 was probably caused by its two pK_a values of 8.6 and 9.9, respectively. However, the change from pH 7 to 2 could not be explained by ionization, and was hypothesized to be due to a switch of rate-determining step in the hydrolysis.

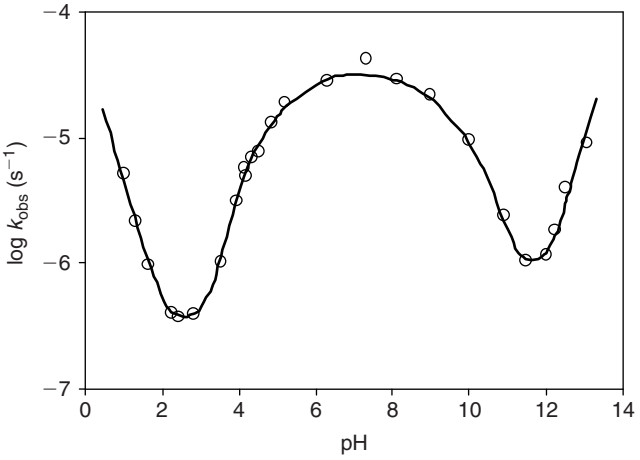


FIGURE 5.9 pH-rate profiles for hydrochlorothiazide at 60°C. *Redrawn from Mollica JA, Rehm CR, Smith JB, Govan HK. Hydrolysis of benzothiadiazines.* J Pharm Sci *1971;**60**:1380–1384.*

5.2.8.4 More complicated pH-rate profiles

The presence of multiple ionization centers can complicate the analysis of a pH-rate profile (Fig. 5.10a). However, their construction is based on similar principles. Depending on how far their pK_a values are isolated, some of the abovementioned features may not be fully developed in a particular pH-rate profile.

Additional features may be identified from a pH-rate profile. The pH-rate profile of aspirin shows evidence for specific acid-catalysis at pH < 2 and specific base-catalysis at pH > 10. The sigmoidal portion (pH 2–9) is due to the different reactivity of the neutral and ionized aspirin species.[7] The broad shoulder at pH 4.5–9 corresponds to the reactivity of ionized aspirin

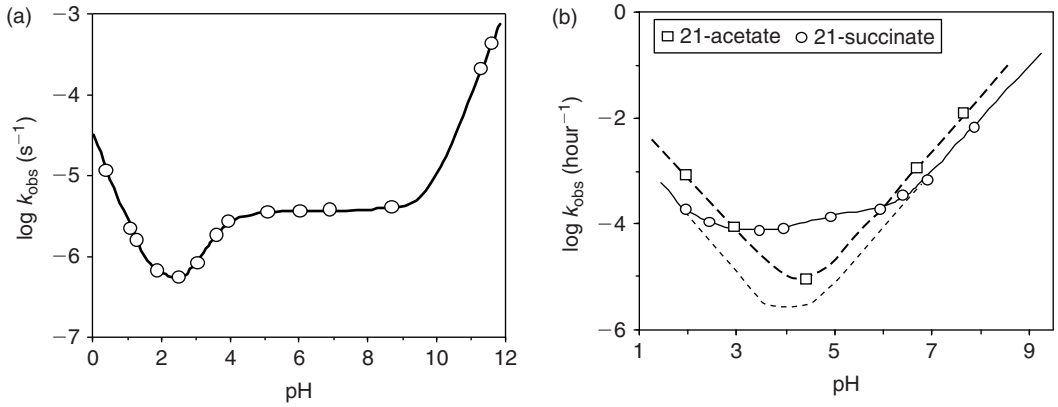
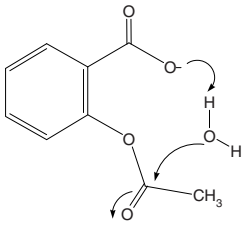
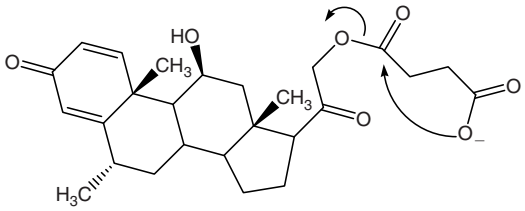


FIGURE 5.10 Complex pH-rate profiles; (a) pH-rate profile for aspirin at 25°C; (b) pH-rate profiles for Methylprednisolone-21-acetate and Methylprednisolone-21-hemisuccinate at 25°C.

species with water, which is due to intramolecular catalysis:



Anderson et al.[21] studied the hydrolysis of methylprednisolone-21-esters in aqueous solutions. Hydrolysis of methylprednisolone-21-acetate (Fig. 5.10b) showed a V-shaped pH-rate profile, indicating both specific acid- and specific base-catalysis. The structurally similar compound, methylprednisolone-21-succinate, however, showed quite a bit of deviation from that expected in the pH range 2–7. The higher reactivity in this pH range and deviation from the V-shaped pH-rate profile was again caused by intramolecular catalysis:



Information like this not only provides significant insights on the stability itself but is also helpful in designing a stable drug candidate.

5.2.9 Solid-state reaction kinetics

Solid dosage forms are the most common because they are easier to handle, more convenient to patients, and do provide stability advantages in most cases. Degradation of an API in solid state is more complicated than in solution, and a large part of its degradation behavior and kinetics can be attributed to the characteristics of its solid form.

5.2.10 Solid-state kinetic models

Various solid-state kinetic models have been discussed in the research literature, most of which have been applied to relatively simple systems, such as API crystals. These reaction models could be useful in elucidating the route and mechanism of solid-state reactions. The degradation kinetics of an API in a solid dosage form is even more complicated. However, zero- or first-order kinetics may be approximated for the first few percentage points of potency loss, which is most relevant for assessing the stability of a pharmaceutical product. Many so-called solid-state reactions might actually be moisture- or solution-mediated.

The concentration term in solid-state reactions is less clearly defined; therefore the extent of conversion, α, is predominantly used. The quantity α is sometimes called the *fractional conversion* or *fraction converted*. Many different reaction mechanisms have been proposed to describe solid-state reactions, some of which are listed in Table 5.2. Models can be expressed in a differential form $f(\alpha_{t,T})$, as well as an integral form $g(\alpha_{t,T})$ or inverse integral form $g^{-1}(k_{t,T})$.

5.2.10.1 *Reactions involving nucleation*

Most solids contain defects, high-energy spots that can initiate physical and chemical changes. These flaws play important roles when nucleation is the rate-limiting step of a reaction.

5.2.10.2 *Avrami–Erofeev equation*

If the reaction involves random nucleation followed by the growth of nuclei of a defined dimensionality,

TABLE 5.2 Solid-State Reaction Mechanisms and Their Corresponding Kinetic Equations

Reaction mechanisms	Differential form $\frac{\partial \alpha}{\partial t} = k_T f(\alpha_{t,T})$	Integral form $g(\alpha_{t,T}) = k_T t$	Inverse integral form $\alpha_{t,T} = g^{-1}(k_T, t)$
REACTION ORDER MODELS			
(F0) Zero-order	1	α	$k_T t$
(F1) First-order (Mampel)	$1-\alpha$	$-\ln(1-\alpha)$	$1-e^{-k_T t}$
(F2) Second-order	$(1-\alpha)^2$	$(1-\alpha)^{-1}-1$	$1-(1+k_T t)^{-1}$
(F3) Third-order	$(1-\alpha)^3$	$\left(\frac{1}{2}\right)((1-\alpha)^{-2}-1)$	$1-(1+2k_T t)^{-1/2}$
NUCLEATION MODELS			
(P2) Power law	$2\alpha^{1/2}$	$\alpha^{1/2}$	$k_T^2 t^2$
(P3) Power law	$3\alpha^{2/3}$	$\alpha^{1/3}$	$k_T^3 t^3$
(P4) Power law	$4\alpha^{3/4}$	$\alpha^{1/4}$	$k_T^4 t^4$
$\left(P\frac{3}{2}\right)$ Power law	$\frac{2}{3}\alpha^{-1/2}$	$\alpha^{3/2}$	$k_T^{2/3} t^{2/3}$
(A2) Avrami–Erofeev	$2(1-\alpha)[-\ln(1-\alpha)]^{1/2}$	$[-\ln(1-\alpha)]^{1/2}$	$1-e^{-k_T^2 t^2}$
(A3) Avrami–Erofeev	$3(1-\alpha)[-\ln(1-\alpha)]^{2/3}$	$[-\ln(1-\alpha)]^{1/3}$	$1-e^{-k_T^3 t^3}$
(A4) Avrami–Erofeev	$4(1-\alpha)[-\ln(1-\alpha)]^{3/4}$	$[-\ln(1-\alpha)]^{1/4}$	$1-e^{-k_T^4 t^4}$
(B1) Prout–Tompkins	$\alpha(1-\alpha)$	$k_T t_{1/2} + \ln[\alpha/(1-\alpha)]$	$(1+e^{-k_T(t-t_{1/2})})^{-1}$
DIFFUSION MODELS			
(D1) One-dimensional	$1/2\alpha$	α^2	$k_T^{1/2} t^{1/2}$
(D2) Two-dimensional	$[-\ln(1-\alpha)]^{-1}$	$[(1-\alpha)\ln(1-\alpha)]+\alpha$	–
(D3) Three-dimensional (Jander)	$\frac{3(1-\alpha)^{1/2}}{2[1-(1-\alpha)^{1/3}]}$	$[1-(1-\alpha)^{1/3}]^2$	$1-(1-k_T^{1/2} t^{1/2})^3$
(D4) Three-dimensional (Ginstling-Brouhnshtein)	$\frac{3}{2}\left[(1-\alpha)^{1/3}-1\right]$	$1-\frac{2\alpha}{3}-(1-\alpha)^{2/3}$	–
GEOMETRICAL CONTRACTION MODELS			
(R1) Linear contraction	1	α	$k_T t$
(R2) Contracting cylinder	$2(1-\alpha)^{1/2}$	$1-(1-\alpha)^{1/2}$	$1-(1-k_T t)^2$
(R3) Contracting sphere	$3(1-\alpha)^{2/3}$	$1-(1-\alpha)^{1/3}$	$1-(1-k_T t)^3$

No simple algebraic solution obtainable; solve by numerical approximation.

the Avrami–Erofeev equation can be derived and may apply[22–26]:

$$[-\ln(1-\alpha)]^{1/n} = kt \tag{5.58}$$

Here, n takes the value of 2, 3, or 4.

Avrami-Erofeev models have been observed for typical crystallization kinetics.

5.2.10.3 *Prout–Tompkins equation*

The Prout–Tompkins equation[27] is derived based on the assumption of linearly growing nuclei that branch into chains:

$$\ln\left(\frac{\alpha}{1-\alpha}\right) = k(t-t_{1/2}) \tag{5.59}$$

Reaction kinetics of the type: solid → solid + gas often follow the Prout-Tompkins model.[28]

5.2.10.4 *Reactions controlled by diffusion*

In some solid-state reactions, reaction rates are controlled by diffusion of the reactants or products. This situation arises because mass transfer in the solid state is usually slow. Examples include curing of polymeric resins.

If the diffusion-controlled process is one-dimensional, then the one-dimensional equation may be derived[29]:

$$\alpha^2 = kt \tag{5.60}$$

Similarly, the rate may be controlled by two-dimensional diffusion, as in a circular disk[30] or in a cylinder, or by three-dimensional diffusion as in a sphere.[31] The following equations can be derived for each case in turn:

$$(1-\alpha)\ln(1-\alpha) + \alpha = kt \tag{5.61}$$

$$1-\frac{2}{3}\alpha-(1-\alpha)^{2/3} = kt \tag{5.62}$$

A simplified three-dimensional diffusion equation was developed by Jander[32]:

$$[1-(1-\alpha)^{1/3}]^2 = kt \tag{5.63}$$

5.2.10.5 Reactions governed by phase boundaries

Some solid-state reactions involve two or more substances, where phase boundaries are formed between the reactants and the products. In some cases, the advance of the phase boundary so formed determines the rate of such a reaction. When the advance of the phase boundary is one-dimensional, then a zero-order reaction is obtained:

$$1-\alpha = kt \tag{5.64}$$

The two-dimensional phase boundary equation (contracting cylinder), and the three-dimensional phase boundary equation (contracting sphere) can be derived accordingly[29]:

$$1-(1-\alpha)^{1/n} = kt \tag{5.65}$$

Here, n takes the value of 2 and 3, respectively.

5.2.10.6 Higher (nth)−order reactions

These kinetic equations have similar formats as in solution kinetics. Most often, the first- and second-order reactions are used. Higher orders are rare and are difficult to interpret in solid-state reactions. In most cases, these equations provide an opportunity for curve-fitting rather than kinetic interpretation at the molecular level.

5.2.10.7 Bawn kinetics

Sometimes the degradant is a liquid that serves as a solvent for the reactant: solid → liquid + gas. As a result, the reactant degrades both in the solid-state and in the solution state. When the kinetics in both phases is first-order, the Bawn kinetic model[28,33] may be appropriate:

$$\ln(1 + B\alpha) = Bk_s t \tag{5.66}$$

Here, k_s is the first-order rate constant in the solid state, and B is the ratio between the first-order rate constant in solution state and that in solid state. Bawn kinetics ceases to apply when the critical conversion is reached, at which time all of the remaining reactant dissolves.

5.2.10.8 Model-fitting versus model-free approaches

Conventionally, kinetic models in solid-state reactions are determined by a model-fitting approach, where the data are fitted to a number of reaction models, and the best model is then selected using statistical criteria. The process is often assisted by other experimental observations, such as microscopy. Difficulties arise for this model-fitting approach because solid-state reactions are usually more complex, and in many cases, a change of reaction mechanism occurs during the course of the reaction. In addition, statistical distinction between models is not always achievable, such as in nonisothermal studies. A different methodology, called the *model-free approach*, has evolved within the last few decades and found application more recently in pharmaceutical systems.[34–36]

The idea of model-free approaches is evident upon transforming the basic kinetic equation:

$$\frac{d\alpha}{dt} = k(T)\,f(\alpha) = A\exp\left(-\frac{E_a}{RT}\right)f(\alpha) \tag{5.67}$$

Here, $d\alpha/dt$ is the instantaneous rate of a reaction, and $k(T)$ is the rate constant at temperature T, which is related to the activation energy E_a and frequency factor A through the Arrhenius relationship. The term $f(\alpha)$ represents the expression of the kinetic model (a function of the extent of conversion). For a reaction studied under a series of different conditions (eg, different isothermal temperatures and different heating rates), the activation energy can be obtained at a specific extent of conversion, α_i, by transforming Eq. (5.67) into

$$\ln\left(\frac{d\alpha}{dt}\right)_{\alpha_i} = \ln[Af(\alpha)]_{\alpha_i} - \frac{E_{\alpha_i}}{RT_{\alpha_i}} \tag{5.68}$$

It is not necessary to assume an explicit form of the reaction model $f(\alpha)$ because at the same conversion, $f(\alpha)$ is a constant number. The Kissenger[37] plot can be considered a special case of model-free kinetics.

In the model-free treatment, the activation energy is obtained as a function of the extent of conversion. If E_a is constant within experimental error with respect to the conversion, the reaction most likely follows a single kinetic model; otherwise, a change of mechanisms is indicated. With this piece of information,

construction of an underlying reaction model can be achieved for reactions following a single mechanism, while a change of mechanism can also be probed for reactions where E_a varies with conversion. Prediction using the model-free approach is also expected to be more accurate for complex solid-state reactions.

5.2.11 Physical parameters affecting solid-state kinetics

Solid-state reactions were hypothesized to consist of four general steps by Paul and Curtin[38]:

1. Molecular loosening at the reaction site
2. Molecular change
3. Solid solution formation
4. Separation of the product phase

Similar steps have been suggested for physical transformations.[39] The four-step hypothesis can provide some insights to help us understand solid-state reactions.

Like reactions in the solution phase, temperature obviously influences the rate of a solid-state reaction, which can be usually be described by the Arrhenius relationship. Due to the constraint of the crystal lattice, solid-state reactions usually require a higher activation energy compared to solution reactions. The applicability of the Arrhenius relationship to solid-state reactions has been debated, but it has generally been observed over narrow temperature ranges. However, a change of temperature beyond that range could induce phase changes, which could alter the validity of the Arrhenius equation. For example, degradation of an amorphous pharmaceutical is not expected to follow the Arrhenius relationship when its glass transition temperature is crossed.[40] In addition, reaction mechanisms could change as temperature is changed.[41] Therefore, caution must be taken when the effect of temperature on solid-state degradation is studied. Some types of reaction are known not to follow the Arrhenius relationship. Free radical oxidative degradation, both in solution and solid states, is such an example, due to the autocatalytic nature of the process. Photolytic reactions are initiated by light absorption and rarely depend on temperature. Indeed, some photochemical reactions can occur at temperatures near absolute zero.

If one envisions the first step loosing of the molecules at the reaction sites, then the site of initiation of a solid-state reaction may have some amorphous characteristics. Hence, the degree of crystallinity of a solid phase is expected to affect the rate of a solid-state reaction. Solid-state reactions are enhanced by crystal disorder, which is related to the fact that amorphous regions have significantly higher molecular mobility, and hence higher reactivity, than crystalline regions. Even when the degree of crystallinity is not altered, a solid-state reaction can be enhanced by increasing its surface area (smaller particle size), and by increasing defects because solid-state reactions are often initiated by defects at the surface of the material. It is not unusual that mechanical treatments such as milling exacerbate a stability problem by increasing the level of defects. However, there are exceptions to this generalization, which will be briefly touched on in the topochemical reactions described next.

Moisture greatly influences solid-state reactions, which will also be discussed separately. Solid-state reactions can also be affected by other factors. For example, drug formulations may decompose faster in a tablet form than in the powder form, primarily due to more intimate contact.

5.2.12 The role of moisture

Water is ubiquitously present as atmospheric moisture and has a profound effect on solid-state reactions. On the one hand, water can act as a reactant and be involved in the reaction itself, such as in hydrolytic reactions. On the other hand, water is an excellent plasticizer; it increases molecular mobility of reactants and enhances drug degradation. In the worst case, water can form a saturated solution phase and make the solid-state reaction solution-mediated, which adversely affects stability to a great degree.

Various mechanisms exist for water molecules to interact with a solid. Water molecules can be incorporated into crystal lattice through hydrogen bonding, van der Waal's interactions, or both. Generally, lattice water is not a cause of chemical instability. Some compounds rely on the interaction of water to form a stable crystal, thus improving their chemical stability. Water can also be adsorbed onto the surface as a monolayer and as multilayers. Water molecules in a monolayer may behave significantly differently than those in the second or third layers, the latter of which have properties that presumably resemble bulk water. The concept of "bound" and "unbound" water has been suggested to explain solid–water interactions; however, it was found to be overly simplistic in some cases. "Bound" water was perceived to be unavailable for drug degradation, while "unbound water" was viewed as the major source of instability of pharmaceuticals. Various models have been applied to describe water–solid sorption, and some success has been achieved in describing the degradation kinetics resulting from such sorption. However, pharmaceutical solids usually contain various defects and, not uncommonly,

various degrees of disordered, amorphous regions. Water molecules are preferentially absorbed into the interior of these regions. Because water is an efficient plasticizer, it significantly increases molecular mobility in the amorphous phase, which causes the drug degradation process to resemble more that occurring in the solution phase. Moisture adsorption onto crystal surfaces only accounts for insignificant amounts (on the order of 0.1% w/w) of total water content. Hence, the moisture level (a few percent) that is usually seen with bulk pharmaceuticals and formulations is primarily from absorption (into the interior), instead of adsorption (on the surface). In addition, at a relative humidity higher than some critical value, the drug or formulations containing the drug may undergo deliquescence, whereby drug degradation occurs primarily in the resulting solution phase (solution-mediated degradation). A more thorough discussion on water–solid interactions in pharmaceutical systems can be found elsewhere.[42]

The influence of moisture on degradation kinetics is complex, and various models have been proposed. The Leeson–Mattocks[43] model hypothesized that a saturated solution layer exists around solid particles, which sheds some light on the importance of moisture. However, very limited successes have been achieved on the quantitative aspect of moisture effects. Genton and Kesselring[44] initially showed the following equation held for nitrazepam degradation in a solid state:

$$\ln k = \ln A + \frac{E_a}{RT + B} \cdot RH \tag{5.69}$$

Therefore, the Arrhenius relationship holds if RH is held constant. Furthermore, the logarithm of the rate constant changes linearly with relative humidity when temperature is held constant. This observation can be understood in the context that the relative humidity only affects molecular mobility in a reaction, but not the reaction pathways. Waterman et al. and Waterman and Adami[45,46] have confirmed this relationship with a number of additional pharmaceutical systems. However, it should be emphasized that this relationship is not universal, and there are many exceptions.

5.2.13 Topochemical reactions

A true solid-state reaction is a topochemical reaction. This type of reaction can be identified if the reaction does not occur in solution, is much slower in solution, or if the reaction products are different than those obtained in solution. True solid-state reactions depend highly on the molecular arrangements (conformations and configurations) or topologies in the solid-state. Therefore, the crystal structure of a solid phase determines its chemical reactivity. Different solid-state forms (polymorphs, solvates, salts, cocrystals) frequently exhibit different reactivity and demonstrate true solid-state reactions.[47]

The first reported example for a topological thermal reaction is the rearrangement of methyl *p*-dimethylaminobenzenesulfonate to the *p*-trimethylammoniumbenzenesulfonate zwitterion.[48,49] The molecules in monoclinic crystals are oriented almost ideally so as to facilitate reactions in the solid state. On the contrary, the reaction rate is much slower (25 times) in its melt than in the solid state. A pharmaceutical example is the deamidation of the asparagine residue of insulin[50]; in this case, the amorphous state is far more stable than the crystalline state. However, a clear interpretation has not been provided, although it is reasonable to believe that the greater reaction rate observed with crystalline insulin should be related to the crystal packing. There are other examples provided in the text by Byrn et al.[51] that different polymorphs lead to different reactivities; they provide other pharmaceutical examples of topochemical reactions.

5.3 COMMON PATHWAYS OF DRUG DEGRADATION

Despite the diversity of drug candidates, drug degradation follows some common pathways. By examining structural features of a drug molecule, possible degradation routes and products may be predicted to a certain extent, which may aid the design and execution of degradation studies.

Common mechanisms of drug degradation include thermolytic, oxidative, and photolytic. The term *thermolytic degradation* refers to those that are driven by heat or high temperature, which normally follows the Arrhenius relationship. Any degradation mechanism can be considered thermolytic if high temperature enhances the rate; in this case, *thermolytic* is being used loosely. Hydrolysis reactions form a subset of thermolytic degradation reactions and are the most common drug degradation pathways; they will be discussed separately. Other thermolytic degradation pathways include ester/amide formation, rearrangement, isomerization/epimerization, cyclization, decarboxylation, hydration/dehydration, dimerization/polymerization, etc. Oxidation of drugs can proceed with different mechanisms: autoxidation, electrophilic/nucleophilic, and electron transfer reactions. Oxidative degradation reactions do not usually follow the Arrhenius relationship due to their complex nature. Photolytic degradation can occur when a drug or drug product is exposed to light. Photolytic degradation is initiated by light absorption; therefore, temperature has a negligible effect. Photolytic degradation is not uncommon,

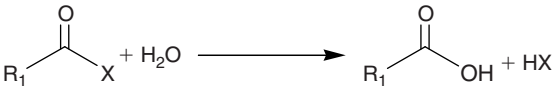
but it may be minimized during the manufacturing, shipping, and storing of drug products in appropriate packaging.

5.3.1 Hydrolysis

Hydrolysis is by far the most commonly encountered drug degradation mechanism, both in solution and in solid state. Many drug molecules contain functional groups derived from relatively weakly bonding groups such as carboxylic acids. Hydrolysis of such derivatives is expected both in solution and in solid state, in the presence of water. In particular, the presence of hydrogen or hydroxyl ions likely catalyzes hydrolytic reactions. In the presence of relevant functional groups, such as alcohols, amines, esters, etc., either as solvent or solvent residuals, or more commonly, as excipients or impurities, the carboxylic acid derivatives can undergo certain types of reactions, such as alcoholysis and transesterification (eg, aspirin-acetaminophen reaction[52]). These reactions follow similar mechanisms to hydrolysis, and will not be treated separately.

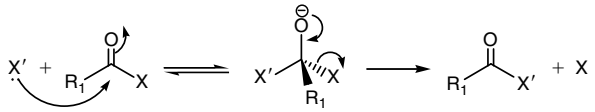
5.3.1.1 *Hydrolysis of carboxylic acid derivatives*

Functional groups derived from carboxylic acid are common in pharmaceuticals. Examples include anhydrides, acyl halides, esters, amides, imides, lactams, lactones, and thiol esters. Esters (including lactones) and amides (including lactams) are among the most commonly seen. Carboxylic acid derivatives are usually prepared by a condensation reaction wherein the carboxylic acid is reacted with the corresponding functional group (alcohol, amine, etc.) with the expulsion of a molecule of water. Since this type of reaction is reversible, the reverse reaction can occur in the presence of added water:



(5.70)

Here X = OCOR, Cl, OR, NR_2, and SR. Sulfonates, sulfonamides, and other sulfonic acid derivatives follow similar pathways. Due to the greater electronegativity of oxygen atoms, the carbonyl group is polarized such that the oxygen atom carries a partial negative charge while the carbon atom carries a partial positive charge. As a result, the carbonyl group can be attacked nucleophilically (by attacking the carbon, using a nucleophile such as OH^-). The hydrolysis of carboxylic acid derivative usually proceeds with a tetrahedral intermediate, as shown here:



(5.71)

In the presence of hydrogen ions, the carbonyl oxygen is protonated, which enhances attack of the carbonyl carbon by weaker nucleophiles such as water. This mechanism explains how the hydrolysis reaction can be catalyzed by either hydrogen or hydroxyl ions (specific acid-base catalysis), as well as by other nucleophiles (general acid-base catalysis).

Insights into the relative hydrolytic stability of carboxylic acid derivatives can be obtained in two aspects: nucleophilic attack of the carbonyl carbon and relative leaving ability of the X group. Hydrolysis will be faster if the nucleophilic attack by X′ is facilitated or if the leaving of X is facilitated, given that other conditions are similar. Electron-withdrawing groups in R_1 increase the electronegativity of the carbonyl group, thus promoting nucleophilic attack on the carbon; electron-withdrawing groups in X increase the basicity of X, stabilize X, and enhance the leaving ability of the X group. On the other hand, electron-donating groups will generally slow down hydrolysis. The effect of electron-withdrawing substituents on R_1 can be seen by comparing the relative hydrolysis rate of a series of substituted benzoate esters, while that of the leaving groups can be seen from the comparison of the hydrolysis rate of different carboxylic derivatives, such as anhydride, ester, amide, and other substances, which usually follows the following pattern: acyl halide > carboxylic anhydride > ester > amide. As a specific example, hydrolysis of phenol esters is much faster than that of corresponding alkyl esters because the phenol anion is a much better leaving group. Because of this, phenol esters are generally more difficult to prepare than the corresponding alkyl esters. More detailed discussion on this topic can be found in a standard organic chemistry textbook, such as Carey and Sundberg.[53]

5.3.1.2 *Hydrolysis of acetals and ketals*

Acetals and ketals are relatively uncommon in pharmaceuticals; however, they are present in some prodrugs. Acetals and ketals are formed by condensation reactions between alcohols and aldehydes and alcohols and ketones, respectively, by removing a water molecule. Due to the unfavorable equilibrium in aqueous solution, and the relative facility of the hydrolysis reaction, they convert back to aldehydes and ketones quickly, particularly in acid solutions:

$$R_2C(OR')_2 + H_2O \xrightarrow{H^+} R_2C{\cdot\cdot}O + 2R'OH \qquad (5.72)$$

The mechanisms of these reactions are the reverse of the formation of the acetals and ketals and are subject to specific acid catalysis, although general acid catalysis is observed in some cases.

5.3.1.3 Hydrolysis of other carbonyl derivatives

A few other types of carbonyl derivatives, such as imines, oximes, hydrazones, and semicarbazones, although not very commonly encountered in drug molecules, can undergo hydrolysis by reversal of the reactions leading to their formation:

$$R_2C{=}NR' + H_2O \rightarrow R_2C{=}O + R'NH_2 \quad (5.73)$$

$$R_2C{=}NOH + H_2O \rightarrow R_2C{=}O + HONH_2 \quad (5.74)$$

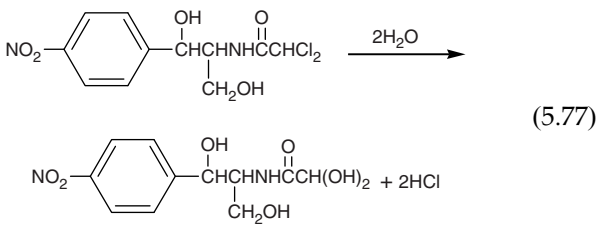
$$R_2C{=}NNHR' + H_2O \rightarrow R_2C{=}O + NH_2NHR' \quad (5.75)$$

$$R_2C{=}NNHC(=O)NH_2 + H_2O \rightarrow R_2C{=}O + NH_2NHC(=O)NH_2 \quad (5.76)$$

Hydrolysis of imines is the most facile of these reactions because the equilibrium constant of formation is unfavorable. On the other hand, oximes, hydrazones, and semicarbazones are usually more stable during hydrolysis because the equilibrium constants for formation are much higher, even in aqueous solution. These hydrolysis reactions are usually catalyzed by both general acids and general bases. Switching of rate-determining steps has also been observed for hydrolysis of imines at different pH values.

5.3.1.4 Miscellaneous hydrolysis reactions

In general, both alkyl halides and aromatic halides are stable to hydrolysis. However, in certain situations, usually when activated by strong electron-withdrawing groups, halides can undergo hydrolysis. An example is chloramphenicol[54]:



(5.77)

The two chlorine atoms are connected to the alpha carbon of the carbonyl, which makes the alpha carbon more nucleophilic, and thus susceptible to hydrolysis.

Ethers hydrolyze to the corresponding alcohols via acidic catalysis. Aryl ethers are particularly unstable under mild acidic conditions, due to the stability of the oxyaryl cation intermediate. Thioethers undergo hydrolysis in a manner similar to ethers. Both ethers and thioethers are relatively stable at neutral or basic pH.

Sulfonamides and sulfonylureas are susceptible to acid hydrolysis. Sulfonamides hydrolyze to the corresponding sulfonic acid and amine, while sulfonylurea decomposes to the amine, carbon dioxide, and sulfonamide. Carbamic esters hydrolyze to the corresponding carbamic acids, which in turn undergo decarboxylation, finally forming the amines. And thiols can hydrolyze to the corresponding alcohols under both acidic and basic conditions.

Epoxides and aziridines are three-member rings, and contain significant strain. Both are susceptible to ring opening reactions via nucleophilic attack by water or other nucleophiles.

5.3.2 Oxidative degradation

Oxidation presents an important drug degradation pathway, second only to hydrolysis. Although its significance to drug stability has been recognized, the study of oxidative stability has not been well developed until more recently, partially due to its complicated nature. This section will briefly discuss the various mechanistic and kinetic aspects of oxidation.

5.3.2.1 Mechanisms of oxidation

Under normal conditions, the reaction between molecular oxygen and an organic molecule, although thermodynamically favorable, proceeds at an insignificant rate. The slow oxidation rate is due to the electronic configuration of molecular oxygen. The valence electrons of the O_2 molecule can be represented, according to molecular orbital theory, as

$$(\sigma_{2s})^2(\sigma^*_{2s})^2(\sigma_{2p_z})^2(\pi_{2p_x})^2(\pi_{2p_y})^2(\pi^*_{2p_y})^1(\pi^*_{2p_z})^1 \quad (5.78)$$

Here, * designates an antibonding orbital. Ground-state molecular oxygen has two unpaired spin electrons in its $\pi^*_{2p_y}$ and $\pi^*_{2p_z}$ orbitals, and therefore it is in a triplet state. Most organic compounds, in contrast, are in a singlet ground state, with all spins paired. Direct reactions between singlet and triplet molecules violate the conservation law of spin angular momentum, and are not favorable. Therefore, for the direct reaction between oxygen and organic molecules to occur, one of the molecules needs to be excited to a state so that the spin states of both molecules match. Triplet ground state oxygen can be excited to the first excited singlet state (singlet oxygen) both chemically and photochemically. Oxidative reactions caused by singlet oxygen will be discussed briefly in the section on photolytic degradation pathways.

Oxidation of organic compounds occurs primarily via one of three mechanisms:

- Nucleophilic/electrophilic processes
- Electron transfer reactions
- Free radical processes (autoxidation)

Nucleophilic/electrophilic processes typically occur between peroxide and an organic reactant. For example, peroxide anion formed under basic conditions can attack drug molecules as a nucleophile. Under normal conditions, a drug will more likely be subject to electrophilic attack by unionized peroxide. Nucleophilic and electrophilic reactions typically follow Arrhenius behavior, and are accelerated at higher temperatures. However, when peroxides are heated, radical processes can be triggered, so non-Arrhenius behavior may be observed above a certain threshold.

In an electron-transfer process, an electron is transferred from a low electron affinity donor to an oxidizing species, which may be catalyzed by transition metals. Single-electron transfer reactions may exhibit Arrhenius behavior when the bond dissociation energy (BDE) of the acceptor molecule is low, but breaking CH bonds requires so much energy that single-electron transfer reactions typically exhibit non-Arrhenius behavior.

The autoxidation process involves the initiation of free radicals, which propagate through reaction with oxygen and drug molecule to form oxidation products. Because of the complexity of the reaction mechanism, non-Arrhenius behavior may be observed. The three stages of autoxidation can be represented as follows:

Initiation:

$$\mathrm{In{-}In} \rightarrow 2\mathrm{In}\bullet \tag{5.79}$$

$$\mathrm{In}\bullet + \mathrm{RH} \rightarrow \mathrm{In{-}H} + \mathrm{R}\bullet \tag{5.80}$$

Propagation:

$$\mathrm{R}\bullet + \mathrm{O_2} \rightarrow \mathrm{ROO}\bullet \tag{5.81}$$

$$\mathrm{ROO}\bullet + \mathrm{RH} \rightarrow \mathrm{ROOH} + \mathrm{R}\bullet \tag{5.82}$$

Termination:

$$\mathrm{R}\bullet + \mathrm{R}\bullet \rightarrow \mathrm{R{-}R} \tag{5.83}$$

$$\mathrm{R}\bullet + \mathrm{ROO}\bullet \rightarrow \mathrm{ROOR} \tag{5.84}$$

$$\mathrm{RO}\bullet + \mathrm{HO}\bullet \rightarrow \mathrm{ROOH} \tag{5.85}$$

In the initiation stage, a pair of free radicals is generated. A variety of processes can generate free radical pairs by cleaving a chemical bond homolytically under the influence of heat or light. Several categories of organic compounds can be used to generate free radicals readily, such as azo compounds, acyl peroxides, alkyl peroxides, and hydrogen peroxide, some of which have been routinely used as free radical initiators in polymerization. For example, the initiation of radicals can be achieved using 2,2′- azobis (*N*,*N*′-dimethyleneisobutyramidine)dihydrochloride (AIBN):

$$(\mathrm{CH_3})_3\mathrm{C{-}N{=}N{-}C}(\mathrm{CH_3})_3 \rightarrow 2(\mathrm{CH_3})_3\mathrm{C}\bullet + \mathrm{N_2} \tag{5.86}$$

The facile dissociation of azo compounds is not caused by the presence of a weak bond, as is the case with peroxides. Instead, the driving force for azo homolysis is the formation of the highly stable nitrogen molecule.

The two steps at the propagation stage are distinct. The first step, the reaction between the drug free radical with oxygen, is fast (on the order of $10^9\ \mathrm{M^{-1}s^{-1}}$ at 300°K).[55] The second step, in which a hydroperoxide is formed and a drug free radical is regenerated, is 8−9 orders of magnitude slower, and it is the rate-determining step. Therefore, the rate of autoxidation is generally unaffected by the oxygen concentration unless the concentration of oxygen becomes significantly low.

The kinetics of autoxidation is complicated due to the multiple steps involved in the process. The reaction order with respect to substrate concentration varies in the range of 0−1. The most noticeable feature of autoxidation kinetics is the lag time, which corresponds to a stage of gradual build up of free radicals.

Besides catalyzing electron transfer processes, transition metals can also react with organic compounds or hydroperoxides to form free radicals:

$$\mathrm{M^{(n+1)+}} + \mathrm{RH} \rightarrow \mathrm{M^{n+}} + \mathrm{R}\bullet + \mathrm{H^+}) \tag{5.87}$$

$$\mathrm{M^{n+}} + \mathrm{ROOH} \rightarrow \mathrm{M^{(n+1)+}} + \mathrm{RO}\bullet + \mathrm{HO^-} \tag{5.88}$$

$$\mathrm{M^{(n+1)+}} + \mathrm{ROOH} \rightarrow \mathrm{M^{n+}} + \mathrm{ROO}\bullet + \mathrm{H^+} \tag{5.89}$$

Typically, multiple mechanisms may occur simultaneously in an oxidative reaction, which introduces further complications.

Stabilization strategies can be developed based on the mechanisms of oxidative degradation for a particular compound.[56,57]

5.3.2.2 *Prediction of oxidative stability*

Oxidative stability of pharmaceuticals can be predicted to a certain extent by theoretical computations.[58] The susceptibility of an organic molecule to nucleophilic/electrophilic and electron transfer oxidative processes can be predicted using frontal molecular orbital (FMO) calculations. According to FMO theory, the site of reaction will occur between the lowest unoccupied molecular orbital (LUMO) of the electrophile and

highest occupied molecular orbital (HOMO) of the nucleophile. Under normal conditions, the nucleophile is usually the organic compound. Therefore, properties of the HOMO of the organic reactant determine its propensity to be oxidized. Computing the HOMO of a reactant could thus yield useful information regarding the most susceptible sites for oxidation by nucleophilic/electrophilic and electron transfer processes.

For oxidative degradation via the free radical process, hydrogen abstraction is usually the rate-determining step for propagation. Therefore, stability of the corresponding free radicals should correlate with their susceptibility to autoxidation. Hence, a BDE calculation could provide insight into sites of potential autoxidation.

5.3.2.3 *Functional groups susceptible to oxidation*

Amines are known to be prone to oxidation. Primary and secondary amines oxidize to hydroxylamines, which can dehydrate to imines or, after further oxidation, to oximes. Oxidation of aryl amines usually produces aryl hydroxylamines, which further oxidize to aryl nitroso compounds. Tertiary amines undergo oxidation, producing *N*-oxides, which are usually the final degradants. Protonation of the amine reduces its propensity to oxidation. However, protonation may not effectively protect against oxidation during long-term storage.

Nitriles are susceptible to oxidation by peroxides under slightly alkaline conditions. Acetonitrile is oxidized to unstable (and oxidizing) peroxycarboximidic acid; this has been known to enhance the susceptibility of organic compounds to oxidation by peroxide when acetonitrile is used as a cosolvent.[59]

Primarily, alcohols can be oxidized to the corresponding aldehydes, and then further to carboxylic acids. Phenols and other aryl hydroxyls are particularly susceptible to oxidation because conjugation in the aryl systems enhances the electron density on the oxygen atoms.

Pyrroles can be oxidized by peroxide to form pyrrolinones. Compounds containing unsaturated double bonds form epoxides by oxidation.

Thiol and thioethers are other examples of readily oxidizable organic functional groups. Thiols can oxidize to disulfide, sulfenic acid, sulfinic acid, and sulfonic acids under a variety of conditions, such as nucleophilic processes (peroxide), autoxidation, and electron-transfer reactions. Due to the affinity of thiols with transition metals, most thiols get oxidized via the metal-catalyzed electron-transfer process. Thioethers undergo oxidation to sulfoxides, and finally to sulfones, a process that also may be catalyzed by transition metals.

A methylene (−CH_2−) or methyne (−CH−group is activated when the corresponding free radical formed by H-abstraction is stabilized by a neighboring group, such as carbonyl, carbon−carbon double bond, aromatic ring, or a heteroatom such as O, N, or S. The activated methylene or methyne group is a potential site of autoxidation, due to the stability of the free radical formed. These sites can be predicted using BDE calculation. Examples include phenylbutazone and benzyl alcohols.

5.3.3 Photochemical degradation

5.3.3.1 *Light*

Light is a form of electromagnetic radiation, the energy of which is given by

$$E = h\nu = \frac{hc}{\lambda} \tag{5.90}$$

Here, h is Plank's constant, c is the speed of light (3×10^8 m/s), ν is the frequency, and λ is the wavelength. Hence, red light (700−800 nm) corresponds to an energy of 30−40 kcal/mol, while far UV light (about 200 nm) corresponds to an energy of 140 kcal/mol. For comparison, the weakest single bonds commonly encountered in organic molecules have a strength of about 35 kcal/mol (such as an O−O bond), and the strongest single bonds have strengths of about 100 kcal/mol (eg, a C−H bond). This simple comparison demonstrates that the energy of light is adequate to potentially lead to degradation of drug molecules if it is absorbed.

No photochemical reaction can occur unless light is absorbed (the Grotthuss−Draper law). The relevant radiation bands that are most likely to be problematic to pharmaceuticals are visible light (λ of about 400−800 nm), and UV light (λ of about 200−400 nm). Sunlight in the wavelength range of 200−290 nm is effectively absorbed by molecular oxygen and ozone in the upper atmosphere, and therefore it is not considered to be important for photolytic degradation of drugs, although many organic substances do absorb strongly in this wavelength range.

5.3.3.2 *Light absorption, excitation, and photochemical reactions*

Accompanying the absorption of UV-visible light by a molecule is a change in its electronic configuration to an excited state. The electrons in the outermost shells of a molecule are the most susceptible ones because these electrons are the least strongly bound to the molecule. From the point of view of FMO theory, electrons in the HOMO will be excited by the absorption of light energy into its LUMO, where they possess higher

potential energy that could cause photochemical degradation of the molecule.

Spin state is a characteristic feature for electron excitation, and it plays an important role in photochemical reactions. Most organic molecules have a ground-state HOMO with paired electrons and are therefore in a singlet state (denoted as S_0, where the subscript 0 refers to the ground state). When one electron is excited from the HOMO to LUMO, its spin can be maintained (first excited singlet state, S_1) or flipped (first excited triplet state, T_1).

The excited electron can drop back to the ground state by giving up one photon; this is called a *photophysical radiative process.* The allowed singlet−singlet ($S_1 \rightarrow S_0$) emission is called *fluorescence,* while the forbidden triplet−singlet ($T_1 \rightarrow S_0$) emission is called *phosphorescence.*

The excited electron can also drop back to the ground electronic state without emitting a photon by photophysical radiationless processes. In most cases, the excited electron falls back to the ground electronic state by releasing heat, which may result from collisions between excited molecules and other molecules (eg, solvent molecules) in their environment. Transitions between states of the same spin (eg, $S_1 \rightarrow S_0 + \text{heat}$) are allowed, and this process is called *internal conversion*; transitions between states of different spins, such as $T_1 \rightarrow S_1 + \text{heat}$ or $T_1 \rightarrow S_0 + \text{heat}$, are forbidden and are called *intersystem crossing.* The energy state diagram shown in Fig. 5.11 describes these photophysical processes.

Not all atoms of a compound absorb UV-visible light, due to the mismatch of the energy of photon and the energetic gap between the ground and excited electronic states. A *chromophore* is defined as an atom or group of atoms that act as a unit in light absorption.

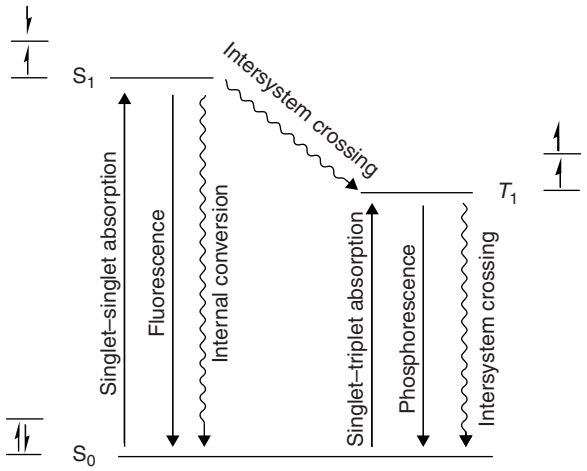


FIGURE 5.11 Schematic energy state diagram for light absorption.

Typical chromophores for organic substances include C:C, C:C, and aromatic groups.

Electronic excitation for C:C bonds involves $\pi \rightarrow \pi^*$ transitions between its electronic orbitals, while for C:C, $n \rightarrow \pi^*$ transitions are involved. Since the electronic occupancy of the excited state is different from that of the ground state, their properties are also different. For example, ethylene is planar in the ground state, while the excited state is twisted. Other properties, such as dipole moments and acid/base dissociation constants, are also changed in the excited state.

Molecules in excited states may undergo chemical reaction with themselves, due to their elevated energy state. The excited molecules may also jump back to ground state via photophysical radiationless routes. The electronic energy is converted to vibrational energy of the molecule in this case, which might be adequate to overcome the activation barrier in the ground state to bond dissociation and trigger chemical degradation.

From the photochemical point of view, most photodegradation reactions observed for pharmaceuticals tend to involve intermediates, and are not concerted in nature. The electronically excited state generated by stretching a σ-bond or twisting a π-bond can result in a diradicaloid structure, where the two half-filled orbitals are nearly degenerate. The interaction of these two degenerated orbitals can then lead to the diradical singlet state, diradical triplet state, and two zwitterionic states (where one degenerated orbital is completely filled while the other is empty). The singlet π, π^* excited state, $S_1(\pi, \pi^*)$, has a significant zwitterionic nature and is expected to form a reactive intermediate via proton or electron transfer reactions, nucleophilic or electrophilic additions, or rearrangement. The triplet $T_1(\pi, \pi^*)$ state, on the other hand, is expected to generate diradical intermediates and undergo primary photoreactions characteristic of radicals (eg, H-abstraction, addition to unsaturated bonds, radicaloid rearrangements, and hemolytic fragmentations). Both $S_1(n, \pi^*)$ and $T_1(n, \pi^*)$ are diradical in nature, and lead to qualitatively similar photochemical products. More detailed discussion of this topic can be found in Turro.[60] Examples of photochemical degradation of drugs can be found in Tønnesen.[61]

5.3.3.3 Photooxidation

Oxidation has frequently been observed in photochemical degradation reactions. Photooxidation can occur when excited triplet state molecules react directly with molecular oxygen because an excited triplet state matches the ground spin state of molecular oxygen. In addition, another process appears to be important in which singlet oxygen is involved, called *photosensitized oxidation.* Photosensitization is a process

whereby an initially light absorbing species (the donor) transfers energy to and excites a second species (the acceptor), provided that the acceptor has a low-lying triplet state. The donor could be a different species (such as a dye molecule), but could also be the drug molecule itself. The acceptor is also called a *quencher* because it quenches the excited state of the donor. Photosensitization allows a nonabsorbing acceptor species to be excited in the presence of a light-absorbing donor molecule.

Two types of photooxidation exist. In Type I photooxidation, the donor transfers an electron or a proton to the acceptor (usually the drug), forming an anion or neutral radical, which rapidly reacts with molecular oxygen. A Type I process is also called an electron transfer or free radical process.

Type II photooxidation involves singlet oxygen. The first excited state of molecular oxygen, $^1\Delta$, lies 22.5 kJ/mol (1270 nm) above its ground state ($^3\Sigma$), which is of much lower energy than the triplet states of most organic molecules. Ground-state molecular oxygen is generally an efficient quencher of the S_1 and T_1 states of organic molecules. The photosensitized singlet oxygen matches the spin states of a majority of organic molecules, which enhances their direct oxidation. Singlet oxygen can be generated using strongly light-absorbing dyes (eg, Rose Bengal, methylene blue) via a photosensitization process.

5.3.4 Other degradation pathways

Hydration (ie, the addition of a molecule of water) can occur at a double bond, which is well known in the case of fumaric acid. Racemization of chiral centers can occur via the formation of a planar intermediate, usually a cation, but a radical or anion intermediate can be involved as well. Decarboxylation often occurs thermally for a β-keto carboxylic acid (eg, norfloxacin[62]).

Amines are common functional groups in APIs, and they are reactive in a variety of ways besides oxidation. Primary and secondary amines can react with counterions in a salt (eg, maleic acid,[63] tartaric acid) or acids/acid derivatives in excipients (eg, stearic acid or magnesium stearate[64]) to form the corresponding amides. Amines react readily with aldehydes (such as formaldehyde, which may present at low levels as a degradation product or impurity in excipients such as polyethylene oxide, polyethylene glycol (PEG), poloxamer, and carbohydrates) to form hemiaminals. The complex series of reactions between a primary or secondary amine and a reducing sugar causes color changes, and is known as the *Maillard browning reaction.*

Other types of degradation reactions have also been observed, such as isomerization (fumaric acid ↔ maleic acid), rearrangements (cephalosporin), dimerization and polymerization (eg, ampicillin and other ß-lactam antibiotics[65]), cyclization (eg, aspartame,[66] moexipril[67]), and deamidation (eg, asparagines residues in peptides and proteins).

5.4 EXPERIMENTAL APPROACHES TO STUDYING THE CHEMICAL DEGRADATION OF DRUGS

Drug degradation studies are a required component of stability testing conducted to support registration of new drug substances and drug products.[68] However, the design and execution of accelerated stability testing methods, while well established in many industries,[69] is still evolving within the pharmaceutical industry. In this section, we survey the current best practices.

5.4.1 Solution thermal degradation studies

In conducting an accelerated thermal degradation study, the investigator typically exposes samples containing the test substance of interest to elevated temperature storage, and then measures the concentration or potency of the test substance remaining after various times. In the past, some methods for the interpretation of data obtained from such studies have focused primarily on the estimation and prediction of reaction rates; however, most users intend to employ these predicted rates to estimate concentrations or potencies at future times at temperatures not tested or to estimate the time until which the concentration or potency will remain within specified limits with a specified probability (shelf life prediction). Although methods that yield estimated rates may be used to obtain point estimates of concentrations, they are incapable of providing useful confidence interval estimates, so this review will focus on methods of obtaining and modeling accelerated degradation experimental data that yield concentration or shelf life estimates directly. Such methods require nonlinear modeling. Because modeling by nonlinear regression is more of an art than a science, previous experience in fitting nonlinear models is helpful.

Chemical kineticists have employed the Arrhenius model for years as a means to determine the energy of activation of a chemical reaction from the temperature dependence of the rate of reaction. Aside from a few studies on biologically derived pharmaceutical preparations,[70] only since the late 1940s did technologists[71–77] in the food and pharmaceutical industries

begin to use the Arrhenius model to study degradation processes. In the mid-1950s, after introductory studies by Garrett, the use of the Arrhenius model to predict the shelf life of a pharmaceutical product at a lower temperature (eg, room temperature) from the kinetics of its degradation studied at higher temperatures became common.[78–86] Prediction of shelf life became a primary focus of thermal stress degradation studies in both the pharmaceutical industry and the food industry, and, over the years, a number of investigators have sought to improve the statistical techniques used to treat the data so as to obtain better estimates of the precision with which shelf life could be predicted. In addition to the use of thermal stress degradation studies to predict shelf life, the time required to determine a reaction rate with a given precision can be reduced by determining the rate at an elevated temperature, where degradation occurs more quickly[87]; thus, Arrhenius models have attracted the interest of chemical engineers and physical chemists.

Early workers in the field of chemical kinetics relied upon graphical methods and simple numerical analysis techniques (such as point-by-point integration) to obtain estimates of the rates of chemical reactions and the effect of changing reaction temperature on these rates. Graphical techniques are important to visualize the data and to form initial judgments as to the overall quality of the data and relevance of the kinetic model; however, they are at best a crude way to obtain kinetic parameters, and are subject to bias on the part of the graphic artist in obtaining the best representation of the data. In addition, graphical methods offer no way to estimate the precision of fitted parameters. Numerical methods based upon point-by-point integration introduce an insidious form of serially correlated error into the final results, as each observation (except the first and last) is used twice. Typical kinetic experiments were run at only a few temperatures, one of which was usually close to ambient. Extraordinary measures to control temperature, and to select very precise analytical methods, are required to employ these methods effectively; in particular, measurement of the energy of activation of a reaction from determination of the rate at only two temperatures about 10°C apart requires temperature control of ± 0.03°C and analytical precision of $\pm 0.1\%$ to obtain an error of $\pm 0.5\%$ in the energy of activation. These methods have been described in detail.[88]

The application of statistical methods, based on modeling by regression analysis to the evaluation of data obtained from accelerated chemical degradation studies in solution, was first applied to a study of the decomposition of a nitrogen-containing compound in 1954 by McBride and Villars.[78] These workers used a two-stage graphical approach, first estimating the rate constants at different temperatures and then plotting the logarithms of the estimated rate constants versus reciprocal absolute temperature in a second step to obtain estimates of the Arrhenius model parameters. Garrett popularized this approach in the pharmaceutical industry in several studies.[77,79,83,84] The use of estimated reaction rates as if they were data reduces the large number of degrees of freedom inherent in the actual data to just a few degrees of freedom in the final model, enormously exaggerating the uncertainty with which interval predictions can be made outside of the time and temperature region studied. Since making such interval predictions is, perhaps, the most important purpose of conducting thermal stress degradation studies, the two-stage modeling approach is wholly unsatisfactory.[89]

Garrett and his followers usually made measurements at linearly spaced time intervals (eg, daily) for a fixed duration (eg, a month or two); measurements at the higher temperatures were terminated when the amount of drug remaining became too low to measure accurately. Most workers also made the same number of measurements at each temperature for each sampling time. This type of design can be called a "Garrett" design. It is not efficient.

Later, Tootill showed that a more efficient design consisted of obtaining measurements at roughly the same extent of decomposition.[90] He was probably the first author to devise a method to use potency remaining at each time and temperature to directly fit an Arrhenius model to kinetic data; however, he did so by imposing strict design constraints. Tootill's 1961 design was actually an extension of the earlier design work of Box and Lucas,[91] who showed that the most efficient design (when percent potency remaining could be measured, the reaction was known to be first order, and the Arrhenius model was known to be correct), was to measure the potency at such a time that 36.8% (100%/e) of the substance remained at both the highest and lowest temperatures that could be studied; this design has properties that statisticians call d-optimal. Davies and coworkers[92,93] showed that the most efficient measurements for the estimation of reaction rates were only obtained after allowing the decomposition reaction to proceed for 1–2 half-lives (50–75% decomposition), in agreement with the finding of Box and Lucas. This recommendation was based solely upon the statistical properties of the data, and ignored potential complications arising from changes in the major reaction pathway caused by the accumulation of degradation products as the reaction proceeds. Thus, Davies and Hudson[93] also recommended designs for which degradation was to be monitored until the same extent of degradation had occurred at each temperature, as previously recommended by Tootill,[90] and by Box and Lucas.[91] Subsequently, such designs will be referred to as "Box-Tootill-Davies" designs. These isoconversional designs are efficient.

Since rates are seldom measured directly, the Arrhenius model can be further integrated to yield, for pseudo-zero-order reactions:

$$P_{hij} = P_0 - t_{ij} A e^{-\frac{E_a}{RT_j}} + e_{hij} \tag{5.91}$$

For pseudo-first-order reactions:

$$P_{hij} = P_0 e^{-t_{ij} A e^{-\frac{E_a}{RT_j}}} + e_{hij} \tag{5.92}$$

Here, P_{hij} is the hth measured potency at time t_{ij} and temperature T_j, P_0 is the estimated initial potency at time $t = 0$, and e_{hij} is the error in measuring P_{hij}. Carstensen and Su introduced this technique in the pharmaceutical field as a method for directly fitting an Arrhenius model to potency data.[94] Nonlinear direct fitting of an Arrhenius model to concentration or potency data was further advanced by the work of Davies and Budgett,[92] who specifically addressed proper weighting of the data. Davies and Hudson later proposed standard methods for the statistical analysis of potency data obtained in the course of degradation studies based on this error analysis.[93] These authors reparameterized the Arrhenius models to explicitly account for the fact that it is the logarithm of estimated rate constants that is approximately normally distributed. Thus, for a pseudo-zero-order reaction:

$$P_{hij} = P_0 - t_{ij} e^{\ln(A) - \frac{E_a}{RT_j}} + e_{hij} \tag{5.93}$$

For a pseudo-first-order reaction:

$$P_{hij} = P_0 e^{-t_{ij} e^{\ln(A) - \frac{E_a}{RT_j}}} + e_{hij} \tag{5.94}$$

Instead of using the logarithm of the rate expected in the limit as the absolute temperature increases to infinity (the Arrhenius frequency factor A), the rate at any other temperature T can be used instead. Thus, one can substitute the following:

$$\ln(A) = \ln(k_{est}) + \frac{E_a}{RT_{est}} \tag{5.95}$$

Furthermore, the rate at any particular temperature is inversely proportional to the time required to achieve a particular fractional degradation f. Thus, for a pseudo-zero-order reaction:

$$k_{T_{est}} = \frac{P_{t=0}(1-f)}{t_{100\ f\%,T_{est}}} \tag{5.96}$$

And for a pseudo-first-order reaction:

$$k_{T_{est}} = \frac{-\ln(f)}{t_{100\ f\%,T_{est}}} \tag{5.97}$$

King et al.[95] independently recommended reparameterization with this additional substitution to estimate shelf life directly. Thus, for a pseudo-zero-order reaction, the Arrhenius model becomes

$$P_{hij} = P_0 - (1-f) P_0 t_{ij} e^{\frac{E_a}{R}\left(\frac{1}{T_{est}} - \frac{1}{T_j}\right) - \ln\left(t_{100\ f\%,T_{est}}\right)} + e_{hij} \tag{5.98}$$

And for a pseudo-first-order reaction, it becomes

$$P_{hij} = P_0 e^{t_{ij}\ \ln(f) e^{\frac{E_a}{R}\left(\frac{1}{T_{est}} - \frac{1}{T_j}\right) - \ln\left(t_{100\ f\%,T_{est}}\right)}} + e_{hij} \tag{5.99}$$

Here, f is the fraction of undegraded drug remaining at time $t_{100f\%}$ at temperature T_{est}.

Simulation studies using a realistic assumed distribution of experimental errors in conjunction with a Box-Tootill-Davies design with $f_{min} = 0.8$ (corresponding to 20% degradation) give remarkably precise estimates of the logarithm of the projected shelf life when either of the abovementioned model equations are used. However, such a design is impractical to execute, since one might need to follow degradation at the lowest temperature included in the accelerated design for a period of time longer than needed to execute a real-time stability study at the temperature of interest T_{est}. Despite this shortcoming, abbreviated studies, in which degradation is not allowed to proceed as far at the lower temperatures, can still give useful results, especially if one's interest is only in projecting whether the projected shelf life at T_{est} is greater than some required minimum (eg, 2 years at $T_{est} = 25°C$).

It is convenient to design accelerated thermal degradation studies so that the temperatures used are selected from a series of equally spaced values on the reciprocal absolute temperature scale (eg, 5°C, 15°C, 25°C, 37°C, 48°C, 61°C, 75°C, 90°C) at temperatures above the final temperature of interest T_{est}. Thus, to project the shelf life at 25°C, one might choose to use the three highest temperatures (eg, 61°C, 75°C, 90°C). If we are only interested in determining whether a drug substance will be stable in solution for 2 years at 25°C, then even this much work is unnecessary. Assuming any reasonable range of expected activation energies, testing at only the two highest temperatures should provide an answer to this question. Of course, the further removed the actual temperatures selected are from the temperature of interest, the larger the error in prediction of the extrapolated value for the logarithm of the shelf life. Nevertheless, reasonable precision is still obtainable using this approach.

It is feasible to design experiments that can be completed in a short period of time (weeks) to determine if the shelf life will be sufficiently long to permit

commercial development of a solution dosage form with storage at either controlled room temperature or under refrigerated storage conditions. All that is needed to design an appropriate experiment is knowledge of the precision achievable by the analytical method that will be used to determine the fraction f remaining undegraded, the allowable extent of degradation f, and the desired storage temperature T_{est}. The design can be validated by simulation prior to execution, and the simulated results can be used to confirm that an adequate number of samples are taken at each temperature. Since most real degradation reactions are pseudo-first-order, sampling times at each temperature should increase exponentially (eg, 0, 1, 2, 4, 7, 14, 28, ... days) so that the change in the fraction remaining will be approximately constant between sampling times. Sufficient samples should be obtained at each temperature to estimate rate constants with equal precision. If one was confident of the order of the reaction, and anticipated degradation rates at each temperature, then the most efficient design would eliminate all intermediate time points, and simply measure the fraction present initially and after some fixed interval storage at each elevated temperature, reducing the analytical workload even further. However, thermal stress degradation studies are usually performed early in the development process, when little is known about the overall stability of the drug substance, so prudent experimenters will want to include samples at intermediate time points to ensure that sufficient usable data is obtained. A pilot experiment, with single samples taken at four temperatures selected from this list on days 0, 1, 2, 4, and 7 (20 samples altogether), will often provide enough information to determine whether additional work is needed. Poorly stable materials can be rejected out of hand, and materials showing no degradation at even the highest temperature require no further study, since no practically achievable activation energy would be consistent with long-term instability if the Arrhenius model holds, and no degradation can be detected after a week of storage at more than 60°C above the intended real-time storage conditions.

Waterman et al. recently extended this concept even further, to incorporate the effects of humidity (Eq. 5.69).[45] These workers have demonstrated that, provided one can design experiments such that the same extent of decomposition is achieved under each of the conditions of temperature and humidity studied (so-called isoconversional conditions), one can estimate both the Arrhenius model parameters and the effect of humidity on the kinetics of decomposition of a drug, either as the API itself or as a component of a formulation, using a small number of measurements with good accuracy. This approach is especially useful for studying solid-state degradation, as isoconversional methods are independent of mechanism.

5.4.2 Solid-state thermal degradation studies

The design of solid-state degradation studies is complicated by the fact that many solid-state degradation reactions do not follow Arrhenius kinetics outside a restricted range. Phase changes (glass transitions, crystalline melting, solvation/desolvation) are expected to radically alter degradation pathways. One must have a thorough understanding of the thermal behavior of the material to be studied prior to undertaking the design of a thermal stress degradation study of a solid material. In most cases, the objective of thermal stress studies on solids is to determine whether moisture will induce hydrolytic degradation. That being the case, the simplest approach is to just add water, stir, and heat.

Isothermal microcalorimetry can be used to study the potential reactivity of solids to water, either alone or in combination with excipients, in relatively short periods of time (days to weeks). Typically, a mixture of the solids and 20% by weight added water is heated to 50−60°C and monitored for thermal reactivity. Hydrolysis reactions are characteristically exothermic and result in the generation of detectable heat signals within the first few days, even for slowly reacting substances. It is prudent to confirm degradation by chemical analysis of any mixture (eg, by high-performance liquid chromatography or liquid chromatography−mass spectrometry) that exhibits the generation of heat. Isothermal microcalorimetry is a completely nonspecific technique that can be used to detect and monitor even the most complex reactions, such as the Maillard reaction between a primary or secondary amine drug and a reducing sugar such as glucose or lactose. This approach is most useful for establishing relative chemical reactivity, such as comparing different potential drug candidates as part of candidate lead selection or for selecting potential excipients to be used in formulation.

A less drastic measure is to store samples of solid materials under conditions of high humidity. Saturated solutions of NaCl maintain essentially 75% relative humidity over a wide range of temperatures, so it is possible to store samples in humidistats (sealed vessels containing a small amount of a saturated solution of NaCl in water) in ovens at different temperatures (eg, 40°C, 60°C, 80°C), provided that no phase changes (such as melting, formation of hydrates, etc.) occur when the samples are heated. Dynamic moisture sorption gravimetric analysis and thermal gravimetric analysis can be used to confirm the absence of hydrate formation at the selected temperatures. Prudent experimenters, not having sufficient test material to perform a thorough characterization of its thermal properties, will opt to begin studies at a lower temperature (eg, 40°C), and progress to studies at higher temperatures

only if it is justifiable to do so when more material becomes available. The very first activity that should be undertaken in the characterization of any new drug substance is to set up a thermal stress degradation experiment for the solid drug substance. This should be started before any other work is undertaken and takes only a few milligrams of material. Samples can be removed from thermal stress conditions at programed times, and stored at low temperature until an analytical method can be developed, assuming larger quantities of material can be obtained later. Often, visual inspection will reveal stability problems, such as color changes and deliquescence, that may dictate storage requirements for the substance being studied. The approach of Waterman and coworkers follows this strategy.[45]

5.4.3 Oxidative degradation studies

Oxidative degradation of drugs is usually caused by impurities (such as peroxide, transition metals, and free radicals) present in the final products. Levels of these impurities are low, variable, and often difficult to control. Hence, oxidation screens are normally conducted in the presence of higher levels of oxidants, catalysts, and other probing reactants. Oxidative testing under these conditions is only prognostic, but not definitive in nature, and quantitative extrapolation to normal storage conditions is challenging and, in most cases, should be avoided. Hydroperoxides, the primary oxidative reagents, may further degrade at elevated temperatures. Therefore, most oxidative tests are carried out at relatively mild temperatures (eg, ambient to 40°C).

As outlined in the section on oxidative degradation, three primary oxidative mechanisms exist: free radical autoxidation, nucleophilic/electrophilic processes, and electron transfer processes. A practical oxidative testing strategy should aim to pinpoint the intrinsic oxidative stability, and the possible oxidative mechanisms. Understanding of oxidative mechanisms allows for stabilization strategies to be developed. Identification of the primary/secondary oxidative degradants also aids in the elucidation of the degradation pathways.

A free radical initiator is frequently used to probe susceptibility of a compound to autoxidation. As mentioned previously, AIBN is a commonly used imitator for this purpose. Its homolytic decomposition half-life in toluene is 10 hours at 65°C, and does not change significantly in different solvents. Tests can be made with molar ratios between AIBN and drug molecule ranging from 0.2 to 1.0. Experiments are often done at 40°C in order to preserve the primary degradants. However, AIBN is not water-soluble. Organic cosolvents, such as acetonitrile and lower alcohols, are necessary for most drug candidates with limited aqueous solubility. Each cosolvent may produce side reactions and has its pros and cons. Water-soluble free radical initiators are available, which have slightly different decomposition kinetics. Examples include 4,4′-azobis(4-cyanopentanoic acid) (ACVA), and 2,2′-azobis(2-amidinopropane) dihydrochloride (AAPH). Antioxidants known to act as free radical trapping agents can also be used to demonstrate radical oxidation in the absence of an added initiator by suppressing oxidation triggered by environmental stress.

Hydrogen peroxide, typically in the concentration range of 0.3−30%, is used to probe the electrophilic/nucleophilic oxidative stability of organic compounds at ambient temperature. Amines, thiols, thioethers, pyrroles, and indoles are functional groups that can be readily oxidized by peroxides. Baertschi[96] stated that, in his experience, a compound is particularly sensitive and could require special efforts for development if a 0.3% solution of hydrogen peroxide induces >20% degradation in 24 hours at room temperature. Alternatively, <5% degradation in 24 hours indicates a relatively stable compound that will likely not present significant challenges to development. Peroxides are present as impurities in many excipients, and some antioxidants are known to be particularly effective as suppressors of peroxide-mediate oxidation. Addition of such antioxidants can help to confirm that peroxide mediated oxidation is responsible for observed degradation.

Many so-called autoxidative degradations may actually be caused by catalysis due to heavy metals, which usually present at very low levels. This is particularly believed for thiols and thioethers, due to their strong interactions with transition metals. In addition, metal ions may react with peroxides, forming reactive hydroxyl radicals. Although levels of heavy metals are low in modern drug substances, their levels may be significant at the early stages. In addition, sensitivity toward metal catalyzed oxidation may be high for certain compounds. Therefore, a transition metal catalysis test still provides valuable information. Compounds that are highly sensitive to these tests may require special processing precautions to avoid metal contamination. Fe^{3+} and Cu^{2+} can be incorporated routinely in oxidative screening at ambient conditions. Chelating agents, such as EDTA, are particularly effective as suppressors of metal-ion catalyzed oxidation. If the addition of chelating agents reduces or prevents oxidative degradation in otherwise untreated samples, metal-ion catalysis due to low levels of contaminating metals is likely to be responsible for the degradation observed.

As an alternative, Tween 80 (or another excipient containing polyoxyethylene moieties), in combination with Fe^{3+} or other transition metals (eg, Mn^{3+}), can be

used to screen the overall oxidative stability of a compound.[96,97] Excipients containing polyoxyethylene moieties, such as Tweens and PEGs, always have low levels of peroxide contaminants. These peroxides react with Fe^{3+}, and produce peroxy radicals that cause the escalation of free radical levels and further oxidation of the excipients. Therefore, the oxidative reaction mechanism in this system could be a combination of all three mechanisms.[97] The Tween test may well be suitable as a first-line oxidative screen. Compounds resisting oxidation in this test could be considered as stable with respect to all three mechanisms of oxidation. For compounds that are susceptible to the Tween test, a further mechanism screen should be carried out using the individual tests.

Susceptibility to singlet oxygen can be tested as a supplement to these oxidative screens, but it is usually not a routine test. A dye (eg, Rose Bengal) can be selected to generate singlet oxygen under light exposure. However, appropriate controls should be applied because many organic compounds are themselves sensitive to photolytic degradation.

In addition to solution tests, a solid-state oxidative screen may provide useful information. Unlike the solution state, reactivity of molecules in the solid state is limited by their crystal packing, and is therefore controlled by their topochemistry. A particular solid form may be stable to oxidative degradation as a solid, even though it is not when dissolved in solution, which can provide criteria to select appropriate polymorphic and salt forms for development. The hydrogen peroxide solvate of urea serves as a convenient source of vapor phase hydrogen peroxide when gently heated, and may be used to supply hydrogen peroxide vapor exposure to solid samples in a manner analogous to the use of saturated salt solutions to supply water vapor at a constant relative humidity.

5.4.4 Photodegradation studies

Photochemical instability is almost never a barrier to successful drug product development. Photochemically labile materials such as sodium nitroprusside are routinely manufactured, distributed, sold, and used to treat patients. The only issue to be resolved experimentally is whether protective packaging may be required, and whether protective lighting systems may be required in laboratory and manufacturing areas.

Photostability testing is required for regulatory approval, and a standard testing protocol for assessing photochemical stability has been established.[98] Pharmacopoeial methods for testing the light resistance properties of proposed packaging materials have also been established.

Commercially available test chambers equipped with an ICH Type I light source, such as a xenon arc lamp, are recommended for initial photochemical stress degradation studies. Chambers equipped with xenon sources typically irradiate samples with more than 2× the required ICH dose of UV radiation when samples are exposed to 1× the required ICH dose of visible light. Unless the test sample absorbs light in the visible region of the electromagnetic energy spectrum, setting the chamber to provide 1× exposure to visible light should provide more than adequate radiation exposure.

In a typical experiment, replicate samples packaged in clear glass containers, clear glass containers wrapped in aluminum foil to exclude all light, and samples in representative colored glass (eg, amber) containers are exposed to the Type I light source following the ICH test conditions. After the exposure has been completed (which may take a few hours, but typically less than 1 day, depending on the intensity of the xenon lamp radiation), the samples are assayed, and any degradation noted. Colored glass containers differ in light transmission properties, so the test should be performed with the least effective (generally the least intensely colored) container being considered. The test may need to be repeated with additional containers having lower specified light transmission.

For the purpose of evaluating laboratory and manufacturing lighting conditions, plastic filter material with defined light transmission properties can be used. A variety of fluorescent lamp tube shields are available commercially that meet this requirement. UV-filter tube shields made of polymethyl methacrylate are available with a cutoff frequency of 395 nm that effectively block all UV light; these are typically sold for use in museums, art galleries, and establishments selling textile products, where fading due to UV light exposure is a potential problem. So-called gold tube shields, designed for use in photolithography manufacturing facilities, effectively block both UV and blue light. Avoid the use of decorative tube shields sold for novelty lighting; many of these transmit some UV light even while blocking almost all blue, green, and yellow visible light. Prepare samples in clear glass containers; place some inside the selected filter tube. Wrap additional samples in aluminum foil to serve as controls. Expose all samples to a Type I light source until the requisite dose of radiation has been achieved. After the exposure has been completed, the samples are assayed. If samples exposed in the clear glass containers show evidence of degradation, but those packaged inside the plastic filter tubes do not, then the plastic filter tubes provide an adequate engineering control that will protect against degradation if installed on lighting fixtures used in production and laboratory areas.

To test for photochemical oxidation, sample containers can be purged with oxygen. Containers purged with an inert gas, such as nitrogen or argon, can serve as controls. A photosensitizing dye, such as Rose Bengal, can be added to test solutions.

Nearly all organic chemicals undergo photochemical degradation if exposed to high-intensity short-wave UV light, such as may be obtained using a high-pressure mercury arc lamp. The degradation products so produced hardly ever reflect those obtained using longer-wave UV or visible light radiation. The use of exotic light sources not likely to be encountered in normal handling and storage is not recommended.

5.5 PHYSICAL STABILITY AND PHASE TRANSFORMATIONS[99]

Physical stability refers to the ability of a solid phase to resist transformation under various conditions. Although in some cases tautomerism may be involved, this section encompasses phase transitions where the chemical entity of APIs remains the same.

As demonstrated in chapter "Crystalline and Amorphous Solids," various solid forms of a chemical entity may have different physicochemical properties. Some of these properties are very relevant to pharmaceutical development, such as solubility/dissolution rate, hygroscopicity, melting point, chemical stability, etc. Extensive efforts are invested early on into the selection of an optimum solid form as an API for downstream development. The understanding of the physical stability of the selected API solid form is minimally necessary to realize the benefit of the selected solid form, to ensure the control during the manufacturing processes, and ultimately to ensure the quality of the finished products.

Transformations among solid forms, mainly as a pure substance, are discussed in detail in chapter "Crystalline and Amorphous Solids," where the applications of these transformations in understanding the interrelationships among the various forms are demonstrated. However, pharmaceutical products consist of many other ingredients, each serving its unique function. Many processing steps are therefore required to bring these ingredients together in a reproducible fashion to achieve certain quality attributes for the finished products. In light of the importance of phase transformations during pharmaceutical processing to preformulation/formulation scientists and processing engineers, the discussions on physical stability in this chapter are biased toward the conditions encountered during pharmaceutical processing.

5.5.1 Types of phase transformations

Various solid forms exist for a certain chemical entity, as categorized in chapter "Crystalline and Amorphous Solids." Many possibilities exist for transformations among these forms. Some phase transformations are thermodynamically favored and are therefore spontaneous. Some phase transformations are thermodynamically disfavored, occur only under stress, and require energy input from the environment. Many phase transformations proceed through a series of individual phase transformations induced by the changes in the environment. Table 5.3 lists the major types of phase transformations.

TABLE 5.3 Major Types of Phase Transformations

Type	Explanation of phase transformation
A	*Polymorphic transition*: Transition between the polymorphs. The crystalline phases include all types of crystalline solids. The composition of the solid remains the same
B	*Hydration/dehydration*: Transition between anhydrates and hydrates or hydrates of different stoichiometry. The compositions of the solids differ by the number of water molecules
C	*Solvation/desolvation*: Transition between solvent-free crystal forms and solvates, solvates of different stoichiometry or solvates of different nature (ie, different solvents are incorporated into the crystalline lattice). The compositions of the solids differ by the nature and number of solvent molecules
D	*Salt/parent conversions or salt/salt exchange*: Transition between the salts (ionic adducts), and the parent unionized compounds (free acids or free bases), between the salts of different stoichiometry or between the different salts. The compositions of the solids differ by the nature and number of counterions
E	*Cocrystal/parent conversions or cocrystal/cocrystal exchange*: Transition between the cocrystals (molecular adducts), and the parent compound (unionized compounds or salts), between the cocrystals of different stoichiometry or between the different cocrystals. The compositions of the solids differ by the nature and number of cocrystal formers
F	*Amorphous crystallization/vitrification*: Transition between crystalline and amorphous phases. The crystalline phases include all types of crystalline solids. Since the compositions of the amorphous phases are usually less well defined, the compositions of the solids change in most cases

5.5.2 Mechanisms of phase transformations

The four underlying mechanisms for phase transformations, and the types of transitions for which these mechanisms could be operating, are listed in Table 5.4. Each of the mechanisms is discussed in further detail.

5.5.2.1 Solid-state transitions

Some phase transitions occur in the solid state without passing through intervening transient liquid or vapor phases. Solid-state reactions have been described previously in Section 5.2.10 and can be mechanistically categorized into four classes: nucleation, geometrical contraction, diffusion, and reaction order models.[100–105] Each model includes several slightly different mechanisms. These mechanisms and their corresponding kinetic equations were summarized in Table 5.2. Recently, a model-free kinetic treatment has been applied to amorphous crystallization,[35] and to dehydration[36] of pharmaceuticals. This new approach is described in detail in Section 5.2.10.6. In summary, this approach provides flexibility in describing the kinetics, and therefore affords better predicting power. Moreover, this approach facilitates the elucidation of the reaction mechanisms, especially those complex ones. In general, the kinetics of phase transition via a solid-state mechanism is influenced by many factors such as the environment (temperature, pressure, relative humidity/partial pressure, etc.), the ideality of the solid phases (the presence and distribution of defects and strains), the physical characteristics of the solid particles (size and distribution, morphology), and the presence of impurities.

5.5.2.2 Melt transitions

When a solid is heated above its melting point and subsequently cooled back to the ambient temperatures, the original solid phase may not be regenerated. Therefore, through this heating/cooling cycle, a phase transition may occur. Among the factors determining the final solid phase are the composition of the melt, relative rates of nucleation, crystal growth, and cooling. Impurities or excipients are also likely to affect the course of crystallization, and thereby the phase transformation. Formulations consist of many components, including API and excipients of different functions. Therefore, it is not only the melting of the API, but also the melting of the excipients that could give rise to phase transitions. In such multicomponent and multiphase systems, it is important to keep in mind phenomena such as melting point depression, eutectic formation among the various components, and the decrease in crystallinity (therefore the melting point) by some upstream pharmaceutical processing.

TABLE 5.4 Underlying Mechanisms of Phase Transformations[99]

Mechanism	Types of phase transformations[a]
Solid-state	A – F
Melt	A, B and C (dehydration/desolvation only), D, E, F (vitrification only)
Solution	A – F
Solution-mediated	A – E, F (amorphous crystallization only); transitions occur only from the metastable phases to the stable phases under the defined conditions

[a]*See Table 5.3 for definitions of types of phase transformations.*

5.5.2.3 Solution transitions

Very often, the solid API will be dissolved, or partially dissolved, in a solvent (typically water) during processing. If subsequent solvent removal induces a transformation, this transformation mechanism is considered a solution mechanism. It is important to note that the transition can be from a metastable phase to the stable phases or vice versa. For instance, the drug may partially dissolve in water during wet granulation or it may completely dissolve in water during freeze-drying or spray drying. Once the solvent is removed, the solid drug will be regenerated from the solution. The regenerated solid may not be the same crystal form as the original phase, and it may consist of a mixture of phases. Thus, through solvent removal, a phase transition may occur. It is also important to realize that only the fraction of drug that is dissolved is capable of undergoing transformation through this mechanism.

The final solid may be a single phase or a mixture of amorphous and crystal forms, depending on the rate of solvent removal, the ease of nucleation and crystal growth of the possible crystal forms under the processing conditions, and the presence of other material in the solution. The undissolved solid drug may serve as seeds and direct the crystallization toward the original crystal form. It may also, together with the insoluble excipients, provide surfaces for heterogeneous nucleation of a different phase. Soluble excipients may have profound effects on the nucleation and crystal growth of the drug, as many crystallization studies have demonstrated.[106]

5.5.2.4 Solution-mediated transitions

As opposed to the solution mechanism, the solution-mediated mechanism only allows the transition from a metastable phase to the stable phases. This type of transformation is driven by the difference in solubility between the two phases. In contrast to the solution mechanism where transformation occurs during drying, the solution-mediated mechanism operates when the metastable phase is in contact with the saturated solution.

Three consecutive steps are involved in a solution-mediated transformation[107–109]:

1. Initial dissolution of the metastable phase into the solution to reach and exceed the solubility of the stable phase
2. Nucleation of the stable phase
3. Crystal growth of the stable phase coupled with the continuous dissolution of the metastable phase

Step 2 or 3 is usually the slowest step. When step 2 is rate-determining, any factor that affects nucleation will influence the overall transformation. These factors include speciation in solution, solubility and solubility difference between the phases, processing temperature, contact surfaces, agitation, and soluble excipients/impurities. When step 3 is the rate-controlling step, the kinetics of the conversion are determined by solubility difference, solid/solvent ratio, agitation, processing temperature, particle size of the original phase, and soluble excipients/impurities.

5.6 PHASE TRANSFORMATIONS DURING PHARMACEUTICAL PROCESSING[99]

5.6.1 Processes for preparing solid dosage forms and associated potential phase transformations

A comprehensive description of the processes used for preparing solid oral dosage forms can be found in many references and textbooks,[110–112] as well as chapters in Section III, design, development and scale-up of formulation and manufacturing process, of this book. Commonly used methods and associated unit operations are summarized in Fig. 5.12. The impact of these processes on solid phase transformations and associated challenges are discussed next.

5.6.1.1 Size reduction

The first step during solid product processing often involves size reduction. Size reduction facilitates subsequent processing and may enhance product performance (eg, through improved morphology/flow properties, minimized segregation, enhanced uniformity, increased surface area, etc.). The principal means for accomplishing size reduction is by milling, which involves shearing/cutting, compressing, and impacting or attrition of drug particles. Impact mills (eg, hammer mills) and fluid-energy mills (eg, jet mills) are widely utilized in the pharmaceutical industry. Other methods for size reduction, such as supercritical fluid technology,[113] are less frequently employed and will not be included in this discussion. Since impact milling typically imparts mechanical stress and often generates heat, it may induce phase transitions, such as polymorphic transitions, dehydration and desolvation, or vitrification via solid-state or melt mechanisms. The rate and extent of these phase transitions will depend on characteristics of the original solid phase, the type of mill, and the milling conditions (such as the energy input). Digoxin, spironolactone, and estradiol are reported to undergo polymorphic transformations during the comminution process.[111] In most cases, the drug substance is milled prior to being mixed with excipients; consequently, the nature and extent of possible phase transformations can be detected more easily relative to the formulation.

5.6.1.2 Granulation/size enlargement

Before a powder can be compressed into a tablet or filled into a capsule, it must possess a number of physical characteristics including flowability, cohesiveness, compressibility, and lubrication. Since most pharmaceutical materials seldom possess all of these properties, granulation methods are frequently used to impart the required characteristics. Wet and dry granulations are the two most commonly used methods of preparing granules for tablet compression or capsule manufacture. Other granulation processes include spray drying and melt granulation, such as high shear melt pelletization, spray-congealing, and melt-extrusion.

5.6.1.2.1 Wet granulation and drying

Wet granulation is most widely used due to its versatility, and the greater probability that the resulting granules will meet the physical requirements for further processing. Wet granulation methods commonly used include low or high shear mixing, fluid-bed mixing, and pelletization (eg, extrusion/spherization).[110] Typical drying methods for the wet granules are tray drying, fluid-bed drying, and vacuum drying. Other methods include microwave drying, tunnel drying, and rotary current drying.[112] Although versatile, wet granulation is especially likely to induce phase transitions. The potential for phase transition of a substance will depend not only on its properties, but also on the conditions and the methods used for granulation and drying.

Conditions such as the amount of liquid used for granulation, the exposure time of the solid to the liquid, airflow, drying temperature, etc., vary with granulation and drying methods. The drug loading, solubility of the drug substance in the granulation liquid, and the procedure for incorporating the drug into the granules will determine whether the starting solid phase is partially or completely dissolved during the granulation process. The subsequent solvent removal rate, which also depends on the method used for granulation and drying, the nature and the concentration of

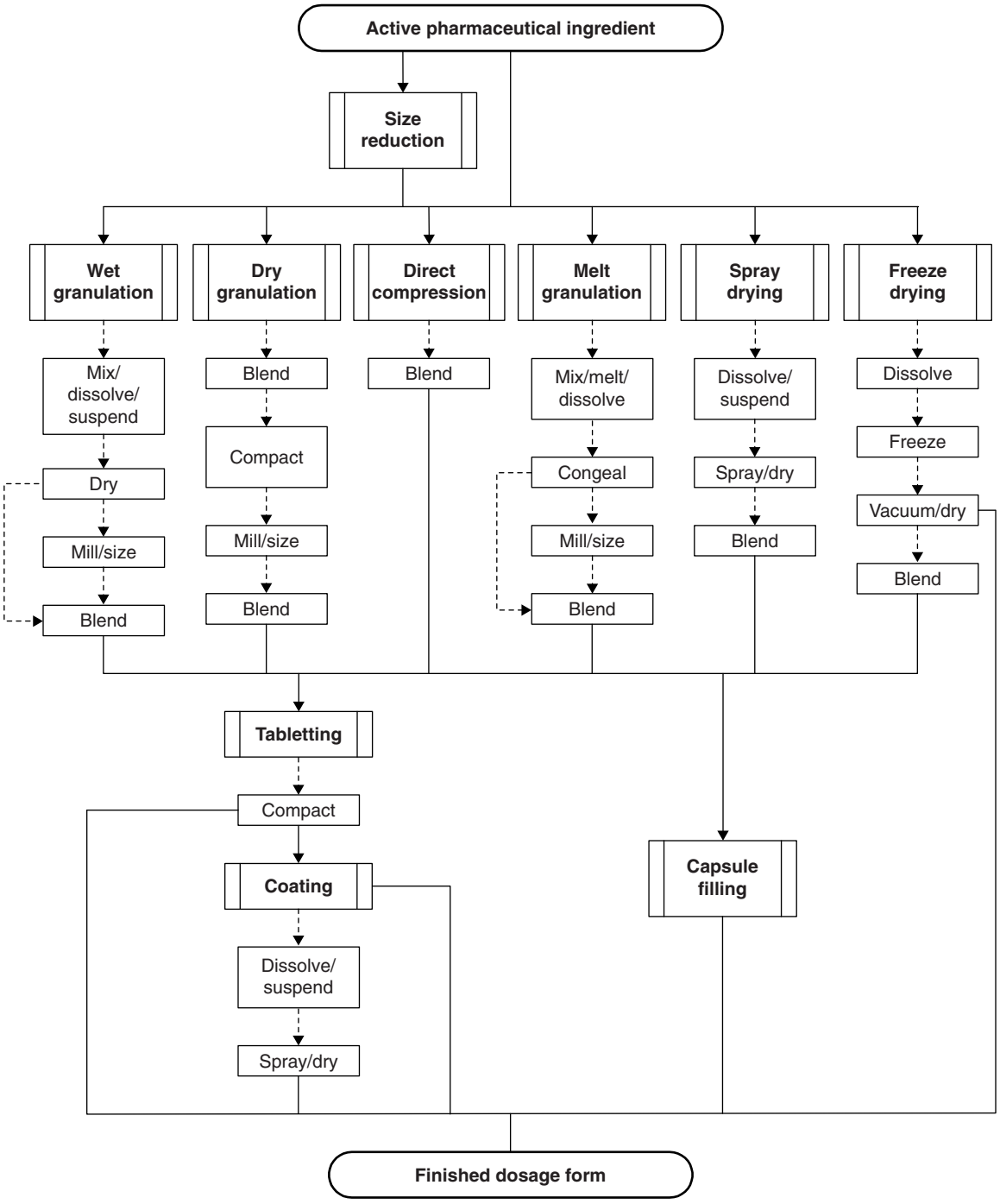


FIGURE 5.12 Common processes for preparing solid oral dosage forms.[99]

the excipient dissolved, may influence the phase formed in the dried granules. Thus, in addition to the properties of the API and the composition of the formulation, the granulating and drying methods and conditions will determine whether solution or solution-mediated phase transformations, such as polymorphic conversion, hydration/dehydration, salt/parent conversion or salt/salt exchange, cocrystal/parent conversion or cocrystal/cocrystal exchange, and vitrification/crystallization will occur. If the drug is completely dissolved in the granulation fluid, the method of drying, as well as the type of excipients used, will determine whether transformation via the solution mechanism will occur. When less soluble drugs are suspended in the granulating fluid, a solution-mediated transformation may occur, leading

to a phase transition from an anhydrous phase to a hydrate, from a salt/cocrystal to its parent, from an amorphous phase to its crystalline counterpart. The ability to detect phase transformations, as well as the number of analytical techniques that can be used for detection, will depend on the drug load in the granules and potential interference from the excipients.

5.6.1.2.2 Dry granulation

Dry granulation is typically used when the formulation ingredients are sensitive to moisture or when they are unable to withstand an elevated drying temperature. It is also used when the formulation ingredients have sufficient inherent binding or cohesive properties. Dry granulation is also referred to as *precompression* or as the *double compression method*.[110] Slugging and roller compaction are the two most commonly used dry granulation methods. The stresses applied during slugging may be lower than those applied by roller compaction because of the differences in dwell time. Since granulating solvent is not used during dry granulation, solution or solution-mediated phase transformations are eliminated, thus the probability of phase transitions with this granulation unit operation is reduced. However, the applied mechanical stresses during processing may lead to phase transformation via the solid-state or melt mechanism. As is the case in the wet granulation process, the ability to detect phase transformations, as well as the number of analytical techniques that can be used, will depend on the drug loading in the granules and on the potential for interference from the excipients.

5.6.1.2.3 Melt granulation

The melt granulation process consists of partially or completely melting solid excipients, and then granulating with the API and other excipients, followed by reducing the mixture to granules by chilling and congealing. During this process, the API is subjected to heat and may be partially or completely dissolved in the molten excipients. If the melting point is relatively low or the heating temperature is sufficiently high, the API may also melt during processing. Partially or completely melted APIs may also serve as binders or congealing carriers. Subsequent cooling could induce phase transitions through the solid-state or melt mechanism. In fact, this process is sometimes used by design for preparing crystalline or amorphous solid dispersions, thus taking advantage of phase transformation through the melt mechanism.

In a spray-congealing process, a low-melting-point carrier is employed to provide the fluidity necessary for spraying. At high temperatures, all or a fraction of the drug may be solubilized in the molten carrier. During spray-congealing, the hot droplets cool rapidly and solidify. Since rapid cooling/congealing is required for the formation of small particles with a narrow size distribution, it is possible that the drug may precipitate as an amorphous phase or as a metastable crystal form, following Ostwald's Rule of Stages.[114,115] In a high-shear-melt granulation, the temperature, extent, and duration of heating, as well as the rate of cooling, are often significantly lower relative to the spray-congealing process. Thus, this process is less likely to initiate phase transitions. Similar to spray-congealing, the melt-extrusion process also requires complete melting of the carrier excipients. Although the rate of cooling is slower, this process is still likely to induce phase transitions.

5.6.1.2.4 Spray drying and freeze-drying

Spray drying produces powder particles that are homogenous, porous, and uniform in size and shape. The technique may also be used for encapsulating or coating drugs to provide protection or release rate control. The process consists of bringing together a highly dispersed liquid and a sufficient volume of hot air to produce uniform droplets. Subsequent evaporation of the liquid leads to particle formation.[110] The feed liquid can be a solution, suspension, or emulsion. This process requires complete or partial dissolution of the drug in a solvent, and thus increases the likelihood for phase transitions involving solution mechanism. Solvent removal from droplets occurs in seconds, and this may lead to rapid crystallization of a metastable phase or the formation of an amorphous phase through the solution mechanism.

Freeze-drying is often used to produce amorphous materials. In this case, extremely low temperatures are used to limit molecular mobility and to prevent nucleation of the drug and the excipients. The conditions for freeze-drying can have a major influence on the solid phase of the drug in the product. Although it is not a commonly used process for manufacturing solid oral dosage forms, freeze-drying has been used to prepare tablets for specific functions (eg, Zydis tablets for rapid dissolution).[116,117]

5.6.1.3 Granulation milling/sizing and blending

A granulated solid is often subjected to milling, and at this stage, the intensity and energy exerted on the API is generally lower than those exerted during initial particle size reduction. Although the risk of phase transition at this stage is less compared to that during API milling and granulation, the presence of excipients makes the detection of phase transitions difficult.

Prior to tablet compression or to encapsulation, granules are directly blended with lubricants and other excipients (eg, lubricants, glidants, and disintegrants) without modifying their physical properties. There is

minimal risk of phase transitions occurring during the blending process.

5.6.1.4 Compression and encapsulation

Lubricated granules are either compressed into tablets or filled into capsules. During tableting, granules may be subject to compression forces as high as 40 kN, with dwell times on the order of a few milliseconds. An energy impact of this magnitude may cause solid phase changes in either the API or the excipients via the solid-state mechanism. For example, caffeine, sulfabenzamide, and maprotiline hydrochloride have been reported to undergo polymorphic transformations during compression.[118] Phase transitions are seldom encountered during the capsule filling process because the solid is experiencing minimal thermal and mechanical perturbations.

5.6.1.5 Coating

When manufacturing finished on the tablet dosage form, a film coating is often applied to each tablet as an aqueous or solvent-based polymer system in coating pans or in a fluid-bed. The function of the coating may be to improve esthetics, mask taste, or modify drug release. Less common coating techniques such as sugar coating, compression coating, or microencapsulation, will not be discussed here. The film coating process involves the application of a thin, polymer-based coating solution to an appropriate substrate (tablets, granules, or crystals) using a spray-atomization technique. To permit uniform distribution of the coating materials and to prevent coating problems (picking, roughness or mottling), film coating parameters are optimized to create a balance between coating solution delivery rate and drying capacity. Rapid drying typically takes place during the application of the film coating. A highly efficient air exchange ensures that there is only a short time between the impingement of the coating liquid onto the tablet surface and the subsequent solvent evaporation. Thus the interaction between the core material and the coating liquid is generally minimal during film coating. In most cases it is unlikely that a phase transition will occur via the solution mechanism during film coating. When necessary, prior to application of the film coat, a polymer-based seal coat may be applied to the surfaces of the tablet cores first. This will prevent extended solid–liquid interactions during the subsequent film coating process.

For some modified release products, a portion of the total dose may be applied as a drug-coating layer. This coat may provide immediate release for fast onset in a biphasic extended release system, or it may provide immediate release in a pulsatile delivery system. The drug layer is typically applied by spraying a drug solution or a drug suspension onto the tablet surface, depending on the drug solubility and dose. Dissolving or suspending the drug in a liquid increases the potential for phase transitions to occur through the solution or solution-mediated mechanism. When a drug solution is used for coating, the probability that such operations will alter the solid phase of the drug is high, similar to that observed in the spray-drying process discussed earlier. That is, rapid solvent removal in this coating process may lead to a transformation to a metastable crystal form or an amorphous phase in the drug layer via the solution mechanism. When a drug suspension is used for coating, if there is a more stable phase in an aqueous medium, the probability that such operations will alter the solid phase of the drug through a solution-mediated mechanism is high. The anticipated risk is higher than that observed in the wet granulation process, and is a result of the following:

- The coating process is usually a much longer process than wet granulation.
- The suspension is usually made with drug particles that are very small (for better suspension, coating uniformity). This leads to faster dissolution to reach supersaturation and, therefore, more time for nucleation of the stable phase to occur.
- The suspension provides a much higher specific solution medium and is usually constantly stirred to ensure the uniformity of the suspension. If nucleation of the stable phase occurs, the conversion to the stable phase will proceed at a much faster rate.

5.6.2 Anticipating and preventing phase transformations in process development

To anticipate and prevent solid phase transitions during manufacturing, it is critical to have a thorough understanding of crystal forms and the amorphous phase of the API and excipients, as well as the interconversion mechanisms and processing options. This integrated knowledge is essential for the rational selection of the physical form of the API, the excipients, the manufacturing process, and for the selection of appropriate handling and storage conditions. In certain cases, even after the solid form and the preferred process are defined, it is advisable to monitor the crystal form of all incoming raw materials and the physical forms present in the final dosage unit. This monitoring is especially important in cases where dissolution or stability of the product is very sensitive to solid phase changes. The rigor used in monitoring will depend on the API, formulation, process, and analytical method. For a highly soluble, stable, and bioavailable molecule, the risk of process-induced phase change on stability and bioavailability may be relatively low. However,

process-induced phase changes in the APIs, excipients, or both may affect manufacturability or disintegration of the dosage form.

In selecting the crystal form of an API for development, the physicochemical, biopharmaceutical, and processing properties must all be taken into consideration. In some cases, it may be necessary to select an alternate crystal form to eliminate stability issues, dissolution rate differences or process-induced phase transitions. It is desirable to choose the crystal form that is least susceptible to phase transformations induced by heat, moisture, and mechanical stresses, provided that the biopharmaceutical and processing characteristics of the API are acceptable. Sometimes an alternate salt having fewer crystal forms may be chosen to minimize process-induced transitions. For example, the authors have encountered a scenario where a new salt form was selected following Phase I studies due to clinical considerations. This change helped us to overcome manufacturing hurdles presented by the complexity of polymorphic phase transitions associated with the original salt form. This change in salt form drastically reduced the number of crystal forms (a single crystal form as opposed to seven processing relevant crystal forms for the original salt, and four processing relevant crystal forms for the parent). This new salt minimized the risk of phase transitions occurring during manufacturing, leading to greater processing flexibility in the development of a solid dosage form for Phase II clinical trials and beyond.

Phase transitions in crystalline excipients, as well as their impact on product performance, also cannot be ignored. For example, process-induced age-hardening in tablets may lead to a decrease in dissolution rates during storage of formulations containing a high level of crystalline excipients such as mannitol. If process-induced hardening is anticipated, variable product dissolution can be minimized through the use of intra- and extra-granular super disintegrants or by selecting alternative excipients.

In designing manufacturing processes for solid dosage forms, process-induced phase transformations can be anticipated based on preformulation studies. These transformations can be controlled and circumvented by selecting the appropriate process. If a solid phase is sensitive to moisture or to solvent, a dry or melt granulation may be used. If a drug substance undergoes an undesirable transition during milling or compression, melt granulation through melt-extrusion may be more desirable provided the drug is thermally stable. It may be possible to avoid milling the drug substance if particle size and shape can be controlled during crystallization. A capsule may be used in place of a tablet dosage form should compression be deemed undesirable. Polymorphic conversion during drying of an enantiotropic polymorph can be avoided by maintaining the drying temperature below the transition temperature. During film coating, solid–liquid interactions at the surface of moisture sensitive cores can be minimized or eliminated by first applying a seal coat that uses a solution of low viscosity at a slow spray rate. Alternatively, an organic solvent-based polymer system can be used for rapid solvent evaporation.

These are just a few examples that illustrate how knowledge of solid-state properties of the APIs and excipients can be applied to formulation design and process selection. Rational formulation and process design can reduce the risk of unpleasant incidents during late stage development and increase the efficiency during new product development. Ultimately, the quality of a solid product manufactured using the chosen process must be confirmed by real-time tests.

References

1. Arrhenius SA. Über die Reaktionsgeschwindigheit bei der Inversion von Rohrzucker durch Saüren. [On the reaction rate for the inversion of sucrose by acids]. *Z Physik Chem* 1889;**4**:226–48.
2. Connors KA, Amidon GL, Stella VJ. *Chemical stability of pharmaceuticals. A handbook for pharmacists*. 2nd ed. New York, NY: John Wiley & Sons, Inc; 1986. p. 19.
3. Connors KA. *Chemical kinetics: the study of reaction rates in solution*. New York, NY: Wiley-VCH Publishers; 1990. p. 191.
4. Eyring H. The activated complex and the absolute rate of chemical reactions. *Chem Rev* 1935;**17**:65–77.
5. Trivedi JS, Porter WR, Fort JJ. Solubility and stability characterization of zileuton in a ternary solvent system. *Eur J Pharm Sci* 1996;**4**:109–16.
6. Alvarez FJ, Slade RT. Kinetics and mechanism of degradation of zileuton, a potent 5-lipoxygenase inhibitor. *Pharm Res* 1992;**9**:1465–73.
7. Garrett ER. Prediction of stability in pharmaceutical preparations. IV. The interdependence of solubility and rate in saturated solutions of acylsalicylates. *J Am Pharm Assoc* 1957;**46**:584–6.
8. Windheuser JJ, Higuchi T. Kinetics of thiamine hydrolysis. *J Pharm Sci* 1962;**51**:354–64.
9. Hou JP, Poole JW. Kinetics and mechanism of degradation of ampicillin in solution. *J Pharm Sci* 1969;**58**:447–54.
10. Zvirblis P, Socholitsky I, Kondritzer AA. The kinetics of the hydrolysis of atropine. *J Am Pharm Assoc* 1956;**45**:450–4.
11. Lund W, Waaler T. The kinetics of atropine and apoatropine in aqueous solutions. *Acta Chem Scand* 1968;**22**:3085–97.
12. Mayer W, Erbe S, Wolf G, Voigt R. Analysis and stability of certain 1,4-benzodiazepines of pharmaceutical interest. 2. Ring contraction in nordiazepam and clonazepam, pH dependence of the hydrolytic cleavage of the diazepam Faustan, and the use of kinetic methods to determine the stability behavior of diazepam in ampul solution. *Die Pharmazie* 1974; **29**:700–7.
13. Yamana T, Tsuji A. Comparative stability of cephalosporins in aqueous solution: kinetics and mechanisms of degradation. *J Pharm Sci* 1976;**65**:1563–74.
14. Hussain A, Schurman P, Peter V, Milosovich G. Kinetics and mechanism of degradation of echothiophate iodide in aqueous solution. *J Pharm Sci* 1968;**57**:411–18.

15. Tabata T, Makino T, Kashihara T, Hirai S, Kitamori N, Toguchi H. Stabilization of a new antiulcer drug (lansoprazole) in the solid dosage forms. *Drug Dev Ind Pharm* 1992;**18**:1437–47.
16. Garrett ER, Nestler HJ, Somodi A. Kinetics and mechanisms of hydrolysis of 5-halouracils. *J Org Chem* 1968;**33**:3460–8.
17. Kondrat'eva AP, Bruns BP, Libinson GS. Stability of D-cycloserine in aqueous solutions at low concentrations. *Khimiko-Farmatsevticheskii Zhurnal* 1971;**5**:38–41.
18. Garrett ER, Bojarski JT, Yakatan GJ. Kinetics of hydrolysis of barbituric acid derivatives. *J Pharm Sci* 1971;**60**:1145–54.
19. Mollica JA, Rehm CR, Smith JB. Hydrolysis of hydrochlorothiazide. *J Pharm Sci* 1969;**58**:635–6.
20. Mollica JA, Rehm CR, Smith JB, Govan HK. Hydrolysis of benzothiadiazines. *J Pharm Sci* 1971;**60**:1380–4.
21. Anderson BD, Conradi RA, Knuth KE. Strategies in the design of solution-stable, water-soluble prodrugs. I: a physical-organic approach to pro-moiety selection for 21-esters of corticosteroids. *J Pharm Sci* 1985;**74**:365–74.
22. Avrami M. Kinetics of phase change. I. General theory. *J Chem Phy* 1939;**7**:1103–12.
23. Avrami M. Kinetics of phase change. II. Transformation-time relations for random distribution of nuclei. *J Chem Phys* 1940;**8**:212–24.
24. Avrami M. Kinetics of phase change. III. Granulation, phase change, and microstructure. *J Chem Phys* 1941;**9**:177–84.
25. Erofeev BV. Generalized equation of chemical kinetics and its application in reactions involving solids. *Compt Rend Acad Sci USSR* 1946;**52**:511–14.
26. Johnson WA, Mehl RF. Reaction kinetics in processes of nucleation and growth. *Trans Am Inst Min Metall Eng* 1939;**135**:416–42.
27. Prout EG, Tompkins FC. Thermal decomposition of KMnO4. *Trans Faraday Soc* 1944;**40**:448–98.
28. Carstensen JT, Rhodes CT. *Drug stability. Principles and practices*. 3rd ed. New York, NY: Marcel Dekker; 2000. p. 154–60.
29. Sharp JH, Brindley GW, Achar BNN. Numerical data for some commonly used solid-state reaction equations. *J Am Ceram Soc* 1966;**49**:379–82.
30. Holt JB, Cutler IB, Wadsworth ME. Rate of thermal dehydration of kaolinite in vacuum. *J Am Ceram Soc* 1962;**45**:133–6.
31. Ginstling AM, Brounshtein BI. The diffusion kinetics of reactions in spherical particles. *J Appl Chem USSR, English Translation* 1950;**23**:1327–38.
32. Jander W. Reaktionen im festen zustande bei höheren temperaturen. I. Mitteilung. Reaktionsgeschwindigkeiten endotherm verlaufender Umsetzungen. *Zeitschrift für Anorganische und Allgemeine Chemie* 1927;**163**:1–30.
33. Bawn CHE. Decomposition of organic solids. In: Garner WE, editor. *Chemistry of the solid state*. London: Butterworths; 1955, Chapter 10.
34. Vyazovkin S. Evaluation of activation energy of thermally stimulated solid-state reactions under arbitrary variation of temperature. *J Comput Chem* 1997;**18**:393–402.
35. Zhou D, Schmitt EA, Zhang GG, Law D, Vyazovkin S, Wight CA, et al. Crystallization kinetics of amorphous nifedipine studied by model-fitting and model-free approaches. *J Pharm Sci* 2003;**92**:1779–92.
36. Zhou D, Schmitt EA, Zhang GGZ, Law D, Wight CA, Vyazovkin S, et al. Model-free treatment of the dehydration kinetics of nedocromil sodium trihydrate. *J Pharm Sci* 2003;**92**:1367–76.
37. Kissinger HE. Reaction kinetics in differential thermal analysis. *Anal Chem* 1957;**29**:1702–6.
38. Paul PC, Curtin DY. Thermally induced organic reactions in the solid state. *Acc Chem Res* 1973;**6**:217–25.
39. Byrn SR, Pfeiffer RR, Stowell JG. *Solid-state chemistry of drugs*. 2nd ed. West Layfayette, IN: SSCI, Inc.; 1999. p. 259–60.
40. Duddu SP, Monte PRD. Effect of glass transition temperature on the stability of lyophilized formulations containing a chimeric therapeutic monoclonal antibody. *Pharm Res* 1997;**14**:591–5.
41. Olsen BA, Perry FM, Snorek SV, Lewellen PL. Accelerated conditions for stability assessment of bulk and formulated cefaclor monohydrate. *Pharm Dev Technol* 1997;**2**:303–12.
42. Zografi G. States of water associated with solids. *Drug Dev Ind Pharm* 1988;**14**:1905–26.
43. Leeson LJ, Mattocks AM. Decomposition of aspirin in the solid state. *J Am Pharm Assoc* 1958;**47**:329–33.
44. Genton D, Kesselring UW. Effect of temperature and relative humidity on nitrazepam stability in solid state. *J Pharm Sci* 1977;**66**:676–80.
45. Waterman KC, Carella AJ, Gumkowski MJ, Lukulay P, MacDonald BC, Roy MC, et al. Improved protocol and data analysis for accelerated shelf-life estimation of solid dosage forms. *Pharm Res* 2007;**24**(4):780–90.
46. Waterman KC, Adami RC. Accelerated aging: prediction of chemical stability of pharmaceuticals. *Int J Pharm* 2005;**293**:101–25.
47. Cohen MD, Green BS. Organic chemistry in the solid state. *Chem Br* 1973;**9**(490–497):517.
48. Sukenik CN, Bonopace JA, Mandel NS, Bergman RC, Lau PY, Wood G. Enhancement of a chemical reaction rate by proper orientation of reacting molecules in the solid state. *J Am Chem Soc* 1975;**97**:5290–1.
49. Sukenik CN, Bonapace JAP, Mandel NS, Lau P-Y, Wood G, Bergman RG. A kinetic and x-ray diffraction study of the solid state rearrangement of methyl *p*-dimethylaminobenzenesulfonate. Reaction rate enhancement due to proper orientation in a crystal. *J Am Chem Soc* 1977;**99**:851–8.
50. Pikal MJ, Rigsbee DR. The stability of insulin in crystalline and amorphous solids: observation of greater stability for the amorphous form. *Pharm Res* 1997;**14**:1379–87.
51. Byrn SR, Pfeiffer RR, Stowell JG. *Solid-state chemistry of drugs*. 2nd ed. West Layfayette, IN: SSCI, Inc.; 1999. p. 333–42.
52. Koshy KT, Lach JL. Stability of aqueous solutions of *N*-acetyl-*p*-aminophenol. *J Pharm Sci* 1961;**50**:113–18.
53. Carey FA, Sundberg RJ. *Advanced organic chemistry. Part A: structure and mechanisms*. 3rd ed. New York, NY: Plenum Press; 1990.
54. Connors KA, Amidon GL, Stella VJ. *Chemical stability of pharmaceuticals. A handbook for pharmacists*. 2nd ed. New York, NY: John Wiley & Sons, Inc; 1986. p. 64.
55. Boccardi G. Oxidative susceptibility testing. In: Baertschi SW, editor. *Pharmaceutical stress testing: predicting drug degradation*, vol. 153. Boca Raton, FL: Taylor & Francis Group; 2005. p. 207.
56. Hovorka SW, Schöneich C. Oxidative degradation of pharmaceuticals: theory, mechanisms and inhibition. *J Pharm Sci* 2001;**90**:253–69.
57. Waterman KC, Adami RC, Alsante KM, Hong J, Landis MS, Lombardo F, et al. Stabilization of pharmaceuticals to oxidative degradation. *Pharm Dev Technol* 2002;**7**:1–32.
58. Reid DL, Calvitt CJ, Zell MT, Miller KG, Kingsmill CA. Early prediction of pharmaceutical oxidation pathways by computational chemistry and forced degradation. *Pharm Res* 2004;**21**:1708–17.
59. Hovorka SW, Hageman MJ, Schöneich C. Oxidative degradation of a sulfonamide-containing 5,6-dihydro-4-hydroxy-2-pyrone in aqueous/organic cosolvent mixtures. *Pharm Res* 2002;**19**:538–45.
60. Turro NJ. *Modern molecular photochemisty*. Sausalito, CA: University Science Books; 1991.
61. Tønnesen HH. *Photostability of drugs and drug formulations*. 2nd ed. Boca Raton, FL: CRC Press; 2004.

62. Florey K. *Analytical profiles of drug substances*. New York, NY: Academic Press; 1991. p. 588.
63. Schildcrout SA, Risley DS, Kleeman RL. Drug-excipient interactions of seproxetine maleate hemi-hydrate: isothermal stress methods. *Drug Dev Ind Pharm* 1993;**19**:1113–30.
64. Florey K. *Analytical profiles of drug substances*. New York, NY: Academic Press; 1991. p. 557.
65. Larsen C, Bundgaard H. Polymerization of penicillins. *J Chromatogr* 1978;**147**:143–50.
66. Leung SS, Grant DJW. Solid state stability studies of model dipeptides: aspartame and aspartylphenylalanine. *J Pharm Sci* 1997;**86**:64–71.
67. Gu L, Strickley RG. Diketopiperazine formation, hydrolysis, and epimerization of the new dipeptide angiotensin-converting enzyme inhibitor RS-10085. *Pharm Res* 1987;**4**:392–7.
68. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH harmonised tripartite guideline Q1A(R2): stability testing of new drug substances and products; 2003.
69. Nelson W. *Accelerated testing: statistical models, test plans and data analysis*. New York, NY: John Wiley & Sons; 1990.
70. Schou SA. Stability of medicaments. *Am J Hosp Pharm* 1960;**17**:5–17.
71. Oswin CR. Kinetics of package life. II. Temperature factor. *J Soc Chem Ind* 1945;**64**:224–5.
72. Lauffer MA. Thermal destruction of influenza-A virus hemagglutinin. II the effect of pH. *Arch Biochem* 1946;**9**:75.
73. Scott EM, Lauffer MA. Thermal destruction of influenza-A virus hemagglutinin. III. The effect of urea. *Arch Biochem* 1946;**11**:179.
74. Brodersen R. Stability of penicillin G in aqueous solution as a function of hydrogen ion concentration and temperature. *Acta Pharmacol Toxicol* 1947;**3**:345.
75. Jellinek H, Urwin JR. The hydrolysis of picolinamide and isonicotinamide in concentrated hydrochloric acid solutions. *J Phys Chem* 1953;**57**:900.
76. Blythe RH. *Shelf life and stability tests in drug packaging*. *Glass Packer*. 1954, August.
77. Garrett ER. Studies on the stability of fumagillin. III. Thermal degradation in the presence and absence of air. *J Am Pharm Assoc Sci Ed* 1954;**43**:539–43.
78. McBride WR, Villars DS. An application of statistics to reaction kinetics. *Anal Chem* 1954;**26**:901–4.
79. Garrett ER, Carper RF. Predictions of stability in pharmaceutical preparations. I. Color stability in a liquid multisulfa preparation. *J Am Pharm Assoc Sci Ed* 1955;**44**:515–18.
80. Levy GB. Accelerated shelf testing. *Drug Cosmet Ind* 1955;**76**:472.
81. Youmans RA, Maasen GC. Correlation of room temperature shelf aging with accelerated aging. *Ind Eng Chem Res* 1955;**47**:1487.
82. Huyberechts S, Halleux A, Kruys P. Une application de calcul statistique à la cinétique chimique. [An application of statistical calculus to chemical kinetics.]. *Bulletin des Societes Chimiques Belges* 1955;**64**:203–9.
83. Garrett ER. Prediction of stability in pharmaceutical preparations. II. Vitamin stability in liquid multivitamin preparations. *J Am Pharm Assoc Sci Ed* 1956;**45**:171–8.
84. Garrett ER. Prediction of stability in pharmaceutical preparations. III. Comparison of vitamin stabilities in different multivitamin preparations. *J Am Pharm Assoc Sci Ed* 1956;**45**:470–3.
85. Lachman L. Prediction of the shelf life of parenteral solutions from accelerated stability studies. *Bull Parenter Drug Assoc* 1959;**13**:8–24.
86. Garrett ER. Prediction of stability of drugs and pharmaceutical preparations. *J Pharm Sci* 1962;**51**:811–33.
87. Haynes JD, Carstensen JT, Callahan JC, Card R. Third Stevens symposium on statistical methods in the chemical industry, 1, cited in Lachman L. Prediction of the shelf life of parenteral solutions from accelerated stability studies. *Bull Parenter Drug Assoc* 1959;**13**:8–24.
88. Benson SW. *The foundations of chemical kinetics*. New York, NY: Mcgraw-Hill Book Co., Inc.; 1960. p. 75–93.
89. Carstensen JT. *Drug stability: principles and practices*. 2nd ed. New York, NY: Marcel Dekker, Inc; 1995. p. 350–9 (revised and expanded).
90. Tootill JPR. Slope-ratio design for accelerated storage tests. *J Pharm Pharmacol* 1961;**13**:75T–86T.
91. Box GEP, Lucas HL. Design of experiments in non-linear situations. *Biometrika* 1959;**46**:77–90.
92. Davies OL, Budgett DA. Accelerated storage tests on pharmaceutical products: effect of error structure of assay and errors in recorded temperature. *J Pharm Pharmacol* 1980;**32**:155–9.
93. Davies OL, Hudson HE. Stability of drugs: accelerated storage tests. In: Buncher CR, Tsay J-Y, editors. *Statistics in the pharmaceutical industry*. New York, NY: Marcel Dekker, Inc; 1981. p. 355–95.
94. Carstensen JT, Su KSE. Statistical aspects of Arrhenius plotting. *Bull Parenter Drug Assoc* 1971;**25**:287–302.
95. King SYP, Kung MS, Fung HL. Statistical prediction of drug stability based on nonlinear parameter estimation. *J Pharm Sci* 1984;**73**:657–62.
96. Baertschi SW. *Pharmaceutical stress testing. Predicting drug degradation*. Boca Raton, FL: Taylor & Francis Group; 2005. p. 43, 63.
97. Harmon PA, Kosuda K, Nelson E, Mowery M, Reed RA. *A novel peroxy radical based oxidative stressing system for ranking the oxidizability of drug substances*. *J Pharm Sci*, 95. 2006, p. 2014–28.
98. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH harmonised tripartite guideline Q1B stability testing: photostability testing of new drug substances and products; 1996.
99. Zhang GGZ, Law D, Schmitt EA, Qui Y. Phase transformation considerations during process development and manufacture of solid oral dosage forms. *Adv Drug Deliv Rev* 2004;**56**:371–90.
100. Garner WE. *Chemistry of the solid state*. New York, NY: Academic Press; 1955.
101. Van Campen L, Monkhouse DC. Solid-state reactions—theoretical and experimental aspects. *Drug Dev Ind Pharm* 1984;**10**:1175–276.
102. Byrn SR, Pfeiffer RR, Stowell JG. *Solid-state chemistry of drugs*. 2nd ed. West Lafayette, IN: SSCI Inc.; 1999.
103. Khawam A, Flanagan DR. *Basics and applications of solid-state kinetics: A pharmaceutical perspective*. *Journal of Pharmaceutical Sciences*, 95. 2006, p. 472–98.
104. Vyazovkin S, Wight CA. Kinetics in solids. *Annu Rev Phys Chem* 1997;**48**:125–49.
105. Brown ME, Dollimore D, Galwey AK. Reactions in the solid state. In: Bamford CH, Tipper CFH, editors. *Comprehensive chemical kinetics*. Amsterdam: Elsevier; 1980. p. 340.
106. Mullin JW. *Crystallization*. 4th ed. Oxford: Elsevier Butterworth-Heinemann; 2001.

107. Cardew PT, Davey RJ. The kinetics of solvent-mediated phase-transformations. *Proc R Soc Lond A Math Phys Eng Sci* 1985;**398**:415–28.
108. Rodríguez-Hornedo N, Lechuga-Ballesteros D, Wu HJ. Phase transition and heterogeneous/epitaxial nucleation of hydrated and anhydrous theophylline crystals. *Int J Pharm* 1992; **85**:149–62.
109. Zhang GGZ, Gu C, Zell MT, Burkhardt RT, Munson EJ, Grant DJW. Crystallization and transitions of sulfamerazine polymorphs. *J Pharm Sci* 2002;**91**:1089–100.
110. Rudnic E, Schwartz JB. *Solid oral dosage forms. Remington's pharmaceutical sciences*. 18th ed. Easton, PA: Mark Publishing; 1990. p. 1633–47.
111. Lieberman HA, Lachman L, Schwartz JB. 2nd ed. *Pharmaceutical dosage forms: tablets*, vol. 1. New York, NY: Marcel Dekker; 1989 (revised and expanded).
112. Lieberman HA, Lachman L, Schwartz JB. 2nd ed. *Pharmaceutical dosage forms: tablets*, vol. 2. New York, NY: Marcel Dekker; 1989 (revised and expanded).
113. York P. Strategies for particle design using supercritical fluid technologies. *Pharm Sci Technol Today* 1999;**2**:430–40.
114. Ostwald W. The formation and changes of solids. *Zeitschrift für Physikalische Chemie* 1897;**22**:289–330.
115. Verma AR, Krishna P. *Polymorphism and polytypism in crystals*. New York, NY: Wiley; 1966. p. 15–30.
116. Seager H. Drug-delivery products and the Zydis fast-dissolving dosage form. *J Pharm Pharmacol* 1998;**50**:375–82.
117. Kearney P. The Zydis oral fast-dissolving dosage form. *Drugs Pharm Sci* 2003;**126**:191–201.
118. Chan HK, Doelker E. Polymorphic transformation of some drugs under compression. *Drug Dev Ind Pharm* 1985; **11**:315–32.

This page intentionally left blank

CHAPTER

6

Excipient Compatibility and Functionality

A.S. Narang, R.V. Mantri and K.S. Raghavan

Drug Product Science & Technology, Bristol-Myers Squibb, New Brunswick, NJ, United States

6.1 INTRODUCTION

Excipients are pharmacologically inert ingredients added intentionally to a drug product (DP) for various functional roles, such as to enhance dosage form volume or size, disintegration of solid dosage forms, binding of particulates, lubrication during processing, taste masking, or modifying drug release. Excipients play a central role in the design and development of dosage forms.[1] The selection of excipients is vital in the design of a quality DP and is based not only on their functionality, but also on the compatibility between the drug and the excipients. Excipient-compatibility testing and formulation development studies are used in the selection of the excipient type, grade, and concentration in formulation. The formulation development and optimization studies also help in identifying the critical material attributes (CMAs) of excipients and active pharmaceutical ingredients (APIs), which is extremely valuable in management of excipient variability and ultimately the robustness of the DP. Excipients are selected to impart a specified function in the oral solid dosage form. Excipients obtained from a wide range of chemical classes may be produced by varied manufacturing processes either in batch or continuous mode, and could be sourced synthetically or from natural products. Excipients can also have multiple functions in the dosage form. These functions are based on the physical and chemical attributes of excipients that impart functionality specific to a dosage form and application. A thorough mechanistic understanding of excipient functionality and its basis is important in deciding on the grade, vendor, and quality attributes (QAs) of selected excipients.

An incompatibility may be defined as an undesirable drug interaction with one or more components of a formulation resulting in changes in physical, chemical, microbiological, or therapeutic properties of the dosage form. Excipient compatibility studies are conducted to mainly predict potential physicochemical incompatibilities of the drug in the final dosage form. These studies also provide justification for selection of excipients and their concentrations in the formulation as required in regulatory filings. Excipient compatibility studies are often thought to be routine and cumbersome. However, these studies are important in the drug development process, as the knowledge gained from excipient compatibility studies is used to select the dosage form components, delineate stability profile of the drug, identify degradation products, and understand the mechanisms of reactions. If the stability of the drug is found to be unsatisfactory in the presence of certain excipients, strategies to mitigate the instability of the drug can be adopted. These strategies can consist of avoidance of a particular excipient, inclusion of a stabilizer (eg, antioxidant for oxidation instability), modifying the process to minimize drug-excipient interaction (eg, extragranular addition of a disintegrant in a wet-granulation formulation), or modifying the kinetics of degradation (eg, by reducing water activity in the dosage form). The excipient compatibility studies are usually carried out in the worst-case scenario of potential proportion of drug-excipient combination in the DP and exposure to the environment (eg, in terms of both temperature and relative humidity (RH)). Thus, kinetic control of incompatibilities to allow an acceptable shelf-life is often feasible.

Methodical and carefully planned and executed compatibility studies can lead to savings in terms of resources and time delays associated with stability issues arising during late-stage product development. The results from these studies can also be useful in determining the causes of stability issues if and when they surface at later stages of development. From a drug development process perspective, these studies are usually conducted after gaining some understanding of solution

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00006-6

and solid-state stability characteristics of the drug substance (API), but before the formulation development activities. An incompatibility in dosage form can result in any of the following: change in color/appearance, loss in mechanical properties (eg, tablet hardness), changes to dissolution performance, physical form conversion, loss through sublimation, decrease in potency, and increase in degradation products.

Additionally, the regulatory expectations have increased significantly over time. This trend is expected to continue as "Quality by Design" initiatives continue to be advocated. Drug-excipient compatibility data is required to justify the selection of formulation components in development reports that go into filings. There has also been an increased regulatory focus on the "Critical Material Attributes" of excipients and their control strategy because of their impact on the DP formulation and manufacturing process. A thorough understanding of the potential variability in the excipient characteristics and attributes that could potentially affect the DP performance is critically important for the lifecycle of the product.

This chapter will focus on the role of excipient functionality, drug-excipient incompatibility, and material (excipient) variability in the design, development, and control of a robust DP.

6.2 EXCIPIENT FUNCTIONALITY

The choice of excipients in the design of a DP is based on several criteria, including salient elements of the quality target product profile, such as the rate and site of drug delivery; stability/target shelf life, packaging, and storage temperature; and appearance, target patient population, and palatability. These combined with input material properties, including the physicochemical and biopharmaceutical properties of the API, and the manufacturing process help determine the selection of excipients. In the process of DP development, QAs of the DP are determined to identify those that may be critical quality attributes (CQAs). Material attributes (MAs) of selected excipients that help realize this goal are enlisted as target material properties. The dependence of CQAs on MAs and manufacturing process parameters helps determine critical material properties (CMPs) and critical product profiles. These, combined with in-process controls and end-product testing protocols, define the control strategy of a DP. During the course of pharmaceutical development, prototype formulations are processed and tested for various DP QAs to finalize the composition of the DP. Once the composition has been identified, aspects such as excipient variability, identification and control of CMAs, definition of DP CQAs, DP specifications, and DP control strategy are put in place to ensure reproducible and robust manufacture of the DP. Identification, assessment, and control of the functionality determining attributes of excipients form an important cornerstone of this process.

Excipients are added to DP formulations based on their function to aid in large-scale manufacture, stability, consistent appearance, patient compliance, and drug delivery and bioavailability. Excipients used in oral solid dosage form have been classified based on their function into groups such as diluents, disintegrants, binders, glidants, lubricants, release-controlling polymers, stabilizers (such as antioxidants, chelators, and pH-modifiers), film-coating polymers, plasticizers, surfactants, colorants, sweeteners, and flavors.

6.2.1 Compendial standards

The reliable use of excipients in dosage form requires reproducible and robust performance in terms of identity, purity, and functionality. The compendial monographs provide specifications that ensure the identity and purity of the excipient, but not necessarily describe the standards for functionality of the excipients. There has been considerable discussion and debate on the excipient functionality, performance, and the role of compendial monographs.[2,3] USP 34 - NF 29 chapter <1059> summarizes a list of several functional categories, along with a general description of mechanisms by which excipients may achieve their function. It also provides a list of physical and chemical properties that could potentially affect the functionality of an excipient.

The European Pharmacopeia (Ph.Eur.) has begun to introduce nonmandatory, functionality-related characteristics (FRCs) in each of the excipient monographs based on the function (or application) to help in identifying the right excipient QA for a specific purpose. The FRCs are defined as *controllable* physical or chemical characteristics of an excipient that affect its functionality, and they may be understood as the control parameters for a specific use of an excipient. The control of excipient FRCs is intended to help provide an assurance of consistent quality and performance of excipients, as well as of the DPs using those excipients.

Table 6.1 provides a limited list of commonly used excipient functionality categories, with commonly used excipients and a summary of recommended properties from the United States Pharmacopeia (USP) and Ph. Eur. compendial agencies as characteristics of excipients that could provide information regarding the functionality of excipients.

6.2.2 Determining FRCs

Excipients can be multifunctional, and the physical and chemical attributes that afford them their functionality

TABLE 6.1 Excipients, Their Functional Categories and Properties/Characteristics That May Potentially Affect Functionality Based on Compendial Guidance

Functional category	Excipient properties that may affect the functional category (extracted from USP 34 - NF 29 chapter <1059>)	Examples of excipients	Functionality-related characteristics (FRCs) for the specific excipient and function (extracted from nonbinding parts of the monographs listed in Ph.Eur.)
Diluent	Particle size and size distribution, particle shape, bulk/tapped/true density, specific surface area, crystallinity, moisture content, powder flow, solubility, compaction properties for tablet dosage forms; chemical properties, composition, including minor concomitant components and undesirable components	Lactose anhydrous	Particle-size distribution Hausner ratio (bulk and tap density) α-lactose and β-lactose: ratio Loss on drying
		Lactose monohydrate	Particle-size distribution Hausner ratio (bulk and tap density)
		Microcrystalline cellulose	Particle-size distribution Powder flow
Binder	Surface tension, particle size, shape and distribution, solubility, viscosity (nature of the polymer structure, molecular weight, and molecular weight distribution) Chemical properties: polymeric structure, monomer properties and sequence, functional groups, degree of substitution, and cross-linking, variations in their sources (for natural polymers)	Povidone	Molecular mass (viscosity, expressed as *K*-value)
		Hydroxypropyl cellulose (HPC)	Viscosity Degree of substitution Powder flow (as matrix polymer) Particle size (as matrix polymer)
		Pregelatinized starch	Cold water-soluble matter Particle-size distribution Powder flow
Disintegrant	Particle size distribution, water absorption rate, swelling ratio or swelling index, the characterization of the resulting product whether it is still particulate or a gel is formed. Chemical composition and influence of dissolution/disintegration pH and ionic nature of APIs (for anionic disintegrants)	Crospovidone	Hydration capacity Particle-size distribution Powder flow Settling volume (for suspensions)
		Croscarmellose sodium	Settling volume Degree of substitution Particle-size distribution Hausner ratio
		Sodium starch glycolate	N/A
Glidant	Particle size distribution, surface area, hygroscopicity, chemical composition	Colloidal silicon dioxide	Specific surface area
Lubricant	Particle size, surface area, hydration state, polymorphic form; solid-state/thermal behavior, purity (eg, stearate: palmitate ratio) moisture content Chemical properties: composition	Stearic acid	Particle-size distribution Specific surface area
		Magnesium stearate	Particle-size distribution Specific surface area Thermogravimetry

differ based on the application. For example, polyvinylpyrrolidone (PVP) may be used as a binder during wet granulation in one application or may be used for stabilization of amorphous API in a solid dispersion application. The physical, chemical, and mechanical attributes of PVP that contribute to its functionality are different for these two applications. For example, hydrophilicity may be counterproductive to the role of PVP in increasing the physical stability of an amorphous drug in the solid dispersion of a poorly soluble drug, whereas this same attribute helps promote particle adhesion when PVP is used as a binder in wet granulation.

Similarly, the functionality determining attributes of hypromellose (hydroxypropylmethyl cellulose, or HPMC) may differ by the application. The Ph.Eur. proposes different FRCs for two different applications of HPMC—namely, (1) viscosity, degree of substitution, molecular mass distribution, particle-size distribution, and powder flow when applied as matrix polymer for extended release dosage forms and (2) only viscosity and degree of substitution when used as a binder, viscosity-increasing agent, or film-former. Similarly, FRCs for the use of povidone as a solubilizer and stabilizer in liquid dosage form include viscosity and molecular mass, while

TABLE 6.2 Potential Impact of Material Properties on Quality Attributes and Processing Behavior[4]

Physical property	Impact Flow	Blending	Wetting	Drying	Mechanical	Dissolution	Stability
Particle size distribution	×	×	×	×	×	×	×
Particle shape distribution	×						
True density				×	×		
Bulk density	×		×		×		
Pore size distribution			×	×		×	
Surface area	×	×	×	×	×	×	×
Surface energy	×	×	×				
Flow	×						
Cohesiveness	×	×					
Internal friction	×				×		
Wall friction	×				×		
Amorphous content			×				×
Elastic modulus					×		
Compactibility					×		
Brittleness					×		
Static charge	×	×					
Hygroscopicity	×			×			×

Reproduced with permission from Hlinak et al. (2006) Understanding critical material properties for solid dosage form design. J Pharm Innov. 1(1) 12–17.

molecular mass has been considered sufficient as an FRC when povidone is used as a binder in tablets and granules.

The compendial monographs can be used as the starting point in developing an understanding of the excipient properties that may be relevant from their functionality perspective, but not as prescriptive tests that define the performance of the excipient in a dosage form. In some cases, the same excipient can affect the performance of the dosage form by multiple functions. For example, microcrystalline cellulose (MCC) is generally used in oral solid formulations as a diluent and to promote disintegration. Nevertheless, there are other characteristics of MCC that contribute to the overall performance of the MCC in the DP, including moisture sorption and flow. The former can contribute to the degradation of moisture-sensitive compounds, while the latter can be particularly important for direct compression formulations. Responding to these needs, the excipient industry routinely provides different particle sizes and moisture grades of MCC with relevant specification controls.

Another example of different MAs affecting DPs through different pathways is evident with the lubricant magnesium stearate (MgSt). While a higher specific surface area of MgSt could be useful in lubrication efficiency, it may also negatively affect disintegration or compaction of the solid dosage form. In many cases, selection of the MgSt quantity in the dosage form and the manufacturing process (number of blending revolutions) needs to balance the adverse impact on compactibility or disintegration with the lubricity needed for flow and to avoid picking and sticking during DP manufacturing. In addition, controls on MgSt specific surface area are often required to provide an assurance of reproducible DP manufacture within a predefined design space.

In a perspective paper, Hlinak et al. (2006) proposed a list of powder properties that could qualify as CMAs, along with their potential impact on product attributes and processing (Table 6.2).[4]

Although this list is not comprehensive, it is instructive to show that many physical attributes of a given material could affect the quality and performance of the DP.

6.2.3 Identification of CMAs

The functional role of excipients in a dosage form is based on the underlying chemistry and the functionality that determines attributes of an excipient. In several cases, chemical attributes such as the degree of substitution on a polymeric excipient affects functionality. For example, the degree of substitution of HPC can affect the wetting of drug particles in a solid dosage form and

drug release or dissolution. The ratio of crystalline α and β forms of lactose could affect the flow and compaction properties. However, it is difficult to specify, a priori, a list of all the functional attributes that will be relevant for the specified dosage form. What is of greater importance for a given product is the identification of critical and performance-indicating attributes of excipients that must be assessed and controlled to achieve a consistent and robust product.

A property of an excipient could be considered to be critical when a realistic change in the excipient property leads to a significant impact on DP performance.[5] Identification of these CMAs is accomplished by developing product and process understanding through systematic studies and extending these studies to study the impact of excipient attributes on the product and the interaction of excipient attributes and the manufacturing process. These studies help in devising an appropriate control strategy to ensure conformance of the DP to prespecified criteria in a quality-by-design (QbD) paradigm. Therefore, the formulation scientist must understand the formulation, manufacturing process, and MAs of the API and excipients before the CQAs are identified and controlled. For example, an oral tablet dosage form of the drug brivanib alaninate (BA) was developed using a wet-granulation process.[6–8] QbD formulation and process robustness studies identified the impact of wet-granulation process parameters on tablet dissolution. The effect of API particle size on tablet dissolution was investigated at the worst-case scenario of the combination of wet-granulation process parameters with respect to tablet dissolution to identify appropriate control limits for the API particle size distribution and specification limits for tablet dissolution. This study highlighted the importance of studying the interaction of process parameters and MAs. In addition, the impact of excipient attributes, such as the degree of substitution of the disintegrant, croscarmellose sodium (CCS), were determined within this design space. The role of the degree of substitution of CCS was identified in affecting drug-excipient binding (Fig. 6.1a), which in turn affects the amount of drug absorbed depending on the drug dose and the drug-CCS binding affinity (Fig. 6.1b). This case study highlights that the degree of substitution of CCS could be a CMA in one scenario, but not in another. The simulations further identify the cases (combination of drug dose and the drug-excipient binding affinity) in which the degree of substitution could be a CQA.

In conducting QbD studies to elucidate the impact of excipient attributes on DP CQAs, it should be recognized that the physicochemical properties of an excipient that might qualify as CMAs may sometimes be different from the properties conventionally characterized or reported on the vendor certificate of analyses or those suggested by the compendial monographs. Risk assessment and prior experience can be utilized to prioritize the MAs that need to be studied during the development studies so that the criticality of these attributes toward the DP quality and performance can be established.

6.3 EXCIPIENT COMPATIBILITY

6.3.1 Chemistry of drug-excipient interactions

The most common reactions observed in pharmaceuticals are hydrolysis, dehydration, isomerization, elimination, cyclization, oxidation, photodegradation, and specific interactions with formulation components (excipients and their impurities). The main factors that affect these reactions are temperature, pH, moisture in solids, RH of the environment, the presence of catalysts, light, oxygen, physical form, and particle size of the drug and excipients.

A comprehensive discussion on chemistry of drug stability is out of the scope of this chapter. However, it is important to consider why excipients may alter the stability of drug substances that are prone to certain types of degradation. Some of the common ways by which excipients may affect drug stability in the dosage form are by altering moisture content in the dosage form, changing microenvironmental pH in the dosage form, acting as general acid/base catalysts, directly reacting with a drug, or becoming the source of impurities that can either react directly with drug substances or participate as catalysts in the drug degradation. The excipients can also alter the physical or the chemical form of the drug through, for example, ion-exchange, transformation of polymorphs, and the formation eutectic or solid solutions. The changes in a physical or chemical state may in turn alter the chemical stability of the drug.

In the context of solid-state compatibility testing, two attributes of excipients are especially important to formulation stability and compatibility testing: (1) their ability to absorb water at variable humidity and (2) the pH that the excipients impart.

6.3.1.1 Influence of water and microenvironmental pH

Most drugs and excipients contain water, which may be either bound or unbound. The bound water is the water of hydration or crystallization, which is so tightly incorporated in the physical form of the material that it is practically immobile and is not available for reactions. This is exemplified by the stability of crystalline hydrates of hydrolytically unstable β-lactam antibiotics, wherein the water is incorporated in the crystalline matrix and is

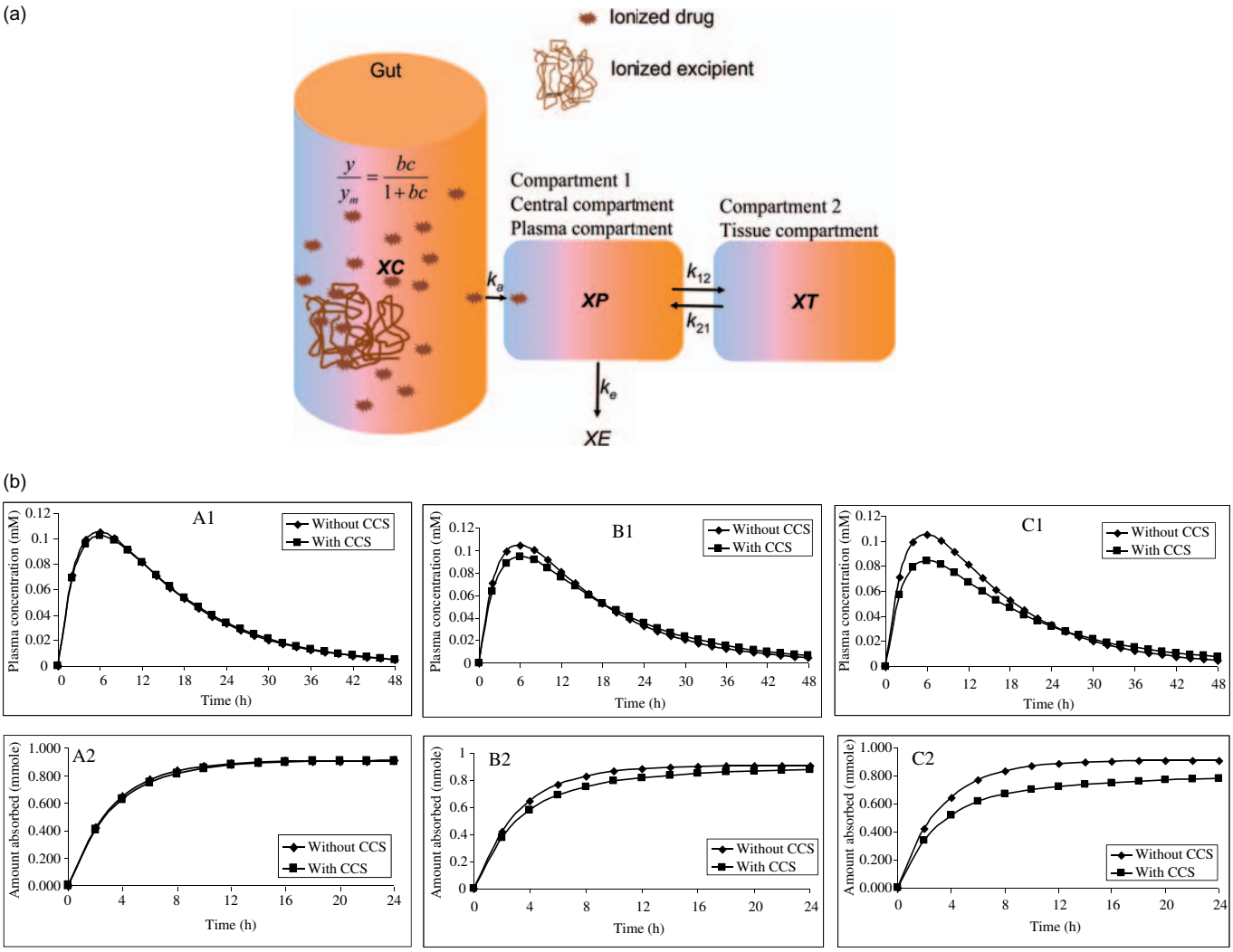


FIGURE 6.1 (a) Pharmacokinetic modeling of the impact of drug-excipient binding on drug absorption from the gastrointestinal (GI) tract; (b) simulation of the effect of drug-excipient binding on oral drug absorption with different degrees of substitution of the excipient, croscarmellose sodium (CCS), affecting pharmacokinetics of the drug, brivanib alaninate (BA), depending on drug dose and binding affinity. The results of simulating the effect of BA-CCS binding on plasma concentration profiles of BA after a single oral dose in the form of a 400-mg tablet using the pharmacokinetic parameters of BA ($V_d = 5.4$ L, $k_a = 0.31$ hour^{-1}, $k_{12} = -0.11$, $k_{21} = 112$, and $k_e = 0.08$) and (A) weak binding ($y_m = 0.5$ mmol BA/mmol CCS and $b = 1$ mM^{-1}), (B) binding parameters observed using Langmuir isotherm ($y_m = 0.85$ mmol BA/mmol CCS and $b = 8.34$ mM^{-1}), or (C) strong binding ($y_m = 2$ mmol BA/mmol CCS and $b = 50$ mM^{-1}). The sub-figures A1, B1, and C1 represent the plasma concentration-time profile. Sub-figures A2, B2, and C2 represent the amount of drug absorbed as a function of time.

not available for reaction. As expected, the stability of these compounds is highly dependent on their crystalline state.[9] In contrast, unbound water usually exists in equilibrium with the atmosphere in an absorbed or adsorbed state by the solid components and has more molecular mobility. The variation in the content of water of an excipient with humidity reflects changes in the unbound water content.

The physical state of water in an excipient or the drug-excipient mixture determines its potential role in drug-excipient interactions. The water sorption-desorption properties of excipients are well documented.[10,11] Presence of water in the solid-state systems has a significant impact on the stability not only in causing the hydrolysis of drugs (eg, of acetylsalicylic acid),[12] but also its role as a reaction medium and in increasing the plasticity and molecular mobility of the system. Excipients that strongly sorb water may prevent drug degradation by scavenging water in a closed system (eg, colloidal silica[13] and silica gel).[14] Excipients with higher adsorption energy can decrease the reactivity of water in the system more than those with lower adsorption energy, as was shown in the case of nitrazepam.[15] On the other hand, water in excipients such as MCC is highly reactive because it is weakly adsorbed.[16]

This was the reason for the higher rate of hydrolytic degradation of aspirin in the presence of MCC versus microfine cellulose.[17] A study of water sorption-desorption of a system as a function of environmental humidity can indicate the strength of sorption of water and its mobility within the system. The mobility of water molecules in a system can be directly measured by nuclear magnetic resonance (NMR) and dielectric relaxation spectroscopy. The mobility of water in the system have been correlated to drug stability in drug-excipient mixtures in several cases; for example, degradation of trichlormethiazide in gelatin gel[18] and of cephalothin in its mixture with MCC.[19]

In addition, water activity is a direct indicator of the amount of free, mobile water in the system. Water activity is proportional to the RH in a closed environment produced in equilibrium with the solid excipient or drug-excipient mixture. Several authors recommend the use of water activity determination for DP stability correlation over the total moisture determination by Karl-Fisher titrimetry, water uptake by weight gain, or both.[20,21] Thus, Burghart et al. correlated the water activity of solid oral dosage forms of levothyroxine and lyothyronine with their chemical stability.[22]

Unbound, weakly adsorbed water contributes to molecular mobility within the system, which is a prerequisite for chemical reactions. Sorbed water plasticizes amorphous solids by reducing the glass transition temperature, Tg.[23,24] The Tg of an amorphous solid represents transition from a highly rigid, less mobile, glassy state to a rubbery, mobile state with higher free volume. Water sorption leading to a reduction in Tg is known in excipients such as starch, lactose, and cellulose,[25] and amorphous drugs such as indomethacin.[26] Molecular mobility of drugs and excipients in the solid state directly correlates with their reactivity.

The pH imparted by the excipient on the microenvironment of the solid state interfaces can have a significant impact on the chemical stability of the drug. For example, the degradation of a fluoropyridinyl drug in a capsule formulation was a function of the pH of the microenvironment facilitating a nucleophilic substitution reaction whereby the fluorine substituent of the pyridine ring was replaced with hydroxyl groups.[27]

Excipients can have an acidic or basic surface pH depending upon their chemical nature and composition. For example, Glombitza et al. measured the surface pH using pH indicator dyes and found that the surface of dicalcium phosphate was more acidic than that of MCC.[28] For soluble excipients, the pH of the excipient solution is a simple indicator of the pH imparted by the excipients in solid state. For insoluble excipients, the pH of 5–20% excipient slurry in water could be used as an indirect indicator. The selection of excipients with compatible pH profiles, based on preformulation solubility and stability studies as a function of pH, is helpful in the design of excipient compatibility experiments. For example, acid labile drugs are expected to be unstable in the presence of acidic excipients such as HPMC phthalate and HPMC-acetate succinate. Similarly, MgSt imparts a basic pH in its microenvironment and may contribute to the instability of base-labile drugs. Stanisz found that the chemical stability of quinapril hydrochloric acid (HCl) in binary drug-excipient mixtures was significantly better with acidic excipients than basic MgSt.[29] This study indicated that both the microenvironmental pH and humidity were significant factors in drug degradation. Thus, the presence of mobile water accelerates the surface pH effects of excipients by creating microenvironmental conditions of dissolved drug on the interacting surfaces.

Most drugs are salts of organic acids or bases, which may exist as free acid or base forms at acidic or basic pH, respectively. Since a minuscule amount of drug may be dissolved in the free water, pH-modifying excipients may result in the formation of the free acid/base form of the drug. If the free acid/base form is more unstable than the salt form, this would lead to enhanced degradation. It may also be volatile and be lost by sublimation from the dosage form, leading to mass balance issues in terms of loss of drug not being accounted for by the presence of degradation products.[30]

6.3.1.2 Reactions with excipients and their impurities

Although generally considered inert, pharmaceutical excipients are organic compounds with functional groups that may undergo chemical reactions in the dosage form, especially with the reactive functional groups of the APIs. Furthermore, pharmaceutical excipients carry trace-level reactive impurities that can either catalyze or directly participate in drug degradation.

A brief review of excipient synthesis, isolation, and purification can give vital clues about their potential impurities and other characteristics that may pose problems in the formulation. However, due to their proprietary nature, the availability of this information is difficult to come by and restricted to informal vendor discussions and perusal of the patent databases in most cases. Several examples of the presence and implication of reactive impurities in pharmaceutical excipients are known in the literature and can be used

as a guiding reference (Table 6.3). Listed here are a few illustrative examples:

- Reactions of reducing sugars (eg, lactose) with primary and secondary amine drugs via Maillard reaction, followed by Amadori rearrangement to produce a multitude of colored products, is well known. The general mechanism of these reactions involves the addition of the amine compound to the open form of the carbohydrate to form an iminium ion intermediate, which can either undergo a ring closure to a glycosamine compound or deprotonate to form the enol version of the Amadori rearrangement product[56] (Fig. 6.2a).
- Even a nonreducing sugar can contain trace levels of reducing sugar. In the case of starch, the terminal glucose was reported to have reacted with hydralazine in the formulation (Fig. 6.2b).[45]
- Formaldehyde and other aldehydes are known impurities in several excipients and packaging components, such as in polyethylene glycol (PEG); see Fig. 6.3. Formaldehyde is known to react with amine drugs to form N-formyl adducts (hemiaminals) that can further react to form dimers (Fig. 6.2c). Adefovir is known to react with formaldehyde to produce the reactive imine, which can further undergo nucleophilic addition with another amine molecule to form a dimer.[21] Nassar et al., showed that BMS-204352 formed an adduct (hemiaminal) with formaldehyde impurity in the solubilizers polysorbate 80 and PEG 300.[22] The impurity of lactose, 5-hydroxymethyl-2-furfuraldehyde, has been reported to react with the carbonyl (ketone) group of haloperidol to form a condensation product (Fig. 6.2d).[31]
- The formyl species could react with amines to form *N*-formamide. For instance, Waterman et al. reported that varenicline, a secondary amine can undergo N-methylation and N-formylation by reactive impurities (formyl and acetyl) found in the polymers used to manufacture osmotic tablets (Fig. 6.2e).[58] Formation of the stearoyl derivative of norfloxacin is attributable to the stearate component of MgSt. Similar to the mechanism shown in Fig. 6.2e, it can be conceived that the secondary amine in norfloxacin would undergo nucleophilic addition to form the stearoyl amide. Seproxetine undergoes a Michael addition reaction with its counterion to form two addition products[59] (Fig. 6.2f).
- The degradation of fosinopril sodium[55] in tablet formulations containing MgSt was attributed to chelation by the magnesium ions.
- Drugs with alcohol groups can form esters with acids (eg, formic acid) or undergo trans-esterification with esters (eg, parabens). Similarly acidic drugs can esterify with excipients containing alcohol groups such as PEGs.
- Trace levels of peroxides and metal ions in formulations are known to accelerate oxidation of drugs. Residual peroxides are also present in PVP, cross-linked PVP, and HPC. Moreover, their content shows both batch-to-batch and manufacturer-to-manufacturer variation.[60] In addition, peroxide levels can change as a function of storage temperature and humidity (Fig. 6.4).[61] The peroxides can cause free radical–initiated oxidation reactions, undergo nucleophilic addition reactions with tertiary amine to *N*-oxide, secondary amine to hydroxylamines, and sulfide to sulfoxide. For example, the hydroperoxide impurity in PVP has been shown to react with a piperazine ring to form N-oxide (Fig. 6.3g).
- Sodium glycolate is a residual reactant in the manufacture of the superdisintegrant sodium starch glycolate, and it has been known to cause degradation of drugs like duloxetine.[62]
- Silicon dioxide may contain significant levels of heavy metal impurities, which may act as catalysts in certain oxidative degradation reactions.

6.3.1.3 *Stabilizing excipients*

Although the focus of compatibility studies is on the identification of potential or seriously destabilizing aspects of drug-excipient interactions, excipients are often also utilized to improve formulation stability. For example, complexation with cyclodextrins has been shown to reduce drug instability in several cases.[63–65] Cyclodextrins internalize the labile drug inside their hydrophobic cavity and shield it from common degradation mechanisms such as hydrolysis, oxidation, and photodegradation.

Stabilization of oxidation-sensitive drugs by the incorporation of antioxidants in the formulation is a well-known strategy. The chosen antioxidants could be water soluble (eg, propyl gallate and ascorbic acid) or water insoluble (eg, butylated hydroxy anisole (BHA), butylated hydroxy toluene (BHT), or α-tocopherol). The choice of antioxidants is based not only on their solubility properties, but also on the mechanism of oxidation. Thus, compatibility studies often involve investigation of the relative efficacy of different antioxidants in mitigating drug degradation.

In addition, the heavy metal–catalyzed degradation reactions in formulations can often be mitigated by the use of chelating agents such as ethylene diamine tetraacetic acid (EDTA) or ethyleneglycol tetraacetic acid (EGTA). For example, EDTA was used to mitigate the oxidative degradation of dextromethorphan in its complexes with the ion exchange resin, divinyl benzene

TABLE 6.3 The Method of Manufacture of Common Pharmaceutical Excipients and Their Potentially Reactive Impurities

Examples of excipients	Method of manufacture	Potentially reactive impurities	Examples of known incompatibilities
Lactose	Lactose is a natural disaccharide consisting of galactose and glucose and is present in the milk of most mammals. Commercially, lactose is produced from the whey of cows' milk, whey being the residual liquid of the milk following cheese and casein production. Cows' milk contains 4.4−5.2% lactose, and it is 38% of the total solid content of milk.[10]	Lactose may contain glucose, furfuraldehyde, formic acid, acetic acid, and potentially other aldehydes.	Maillard reactions, Claissen-Schmidt condensation reaction of its impurity -hydroxymethyl-2-furfuraldehyde,[31] and catalysis of hydrolysis.[32,33]
Microcrystalline cellulose	Microcrystalline cellulose is manufactured by the controlled hydrolysis, with dilute mineral acid solutions of α-cellulose, obtained as a pulp from fibrous plant materials. Following hydrolysis, the hydrocellulose is purified by filtration and the aqueous slurry is spray-dried to form dry, porous particles of a broad-size distribution.[10]	The impurities in microcrystalline cellulose are glucose, formaldehyde, nitrates, and nitrites.	Water sorption resulting in increased hydrolysis,[17] Maillard reaction with residual glucose,[34] adsorption of basic drugs,[35] and nonspecific incompatibilities due to hydrogen-bonding capability.[36]
Povidone and crospovidone	Pyrrolidone is produced by reacting butyrolactone with ammonia. This is followed by a vinylation reaction in which pyrrolidone and acetylene react under pressure. The monomer, vinylpyrrolidone, is then polymerized in the presence of a combination of catalysts to produce povidone. Water-insoluble, cross-linked PVP (crospovidone) is manufactured by a polymerization process where the cross-linking agent is generated in situ.[10]	Povidone and crospovidone contain significant levels of peroxides. Povidone may also contain formic acid and formaldehyde.[37]	Oxidation attributable to peroxides,[38] nucleophilic addition to amino acids and peptides,[39] and hydrolysis of sensitive drugs due to moisture.
Hydroxypropyl cellulose (HPC)	HPC is a water-soluble cellulose ether produced by the reaction of cellulose with propylene oxide.[10]	HPC may contain significant levels of peroxides.	Oxidation of sensitive drugs due to residual peroxides.
Croscarmellose sodium	To produce croscarmellose sodium, alkali cellulose is prepared by steeping cellulose, obtained from wood pulp or cotton fibers, in sodium hydroxide solution. The alkali cellulose then reacts with sodium monochloroacetate to obtain carboxymethylcellulose sodium. After the substitution reaction is completed and all the sodium hydroxide has been used, the excess sodium monochloracetate slowly hydrolyzes to glycolic acid. The glycolic acid changes a few of the sodium carboxymethyl groups to free acid and catalyzes the formation of cross-links to produce croscarmellose sodium. The croscarmellose sodium is then extracted with aqueous alcohol and any remaining sodium chloride or sodium glycolate removed. After purification, croscarmellose sodium of greater than 99.5% purity is obtained. The croscarmellose sodium may be milled to break the polymer fibers into shorter lengths and hence improve its flow properties.[10]	Monochloroacetate, nitriles, and nitrates. Monochloroacetate can react with nucleophiles.	Weakly basic drugs can compete with the sodium counterion, thus getting adsorbed on the surface of the disintegrant particles.[40,41] Drug salt form conversion has also been reported.[42]

(*Continued*)

TABLE 6.3 (Continued)

Examples of excipients	Method of manufacture	Potentially reactive impurities	Examples of known incompatibilities
Sodium starch glycolate	Sodium starch glycolate is a substituted and cross-linked derivative of potato starch. Starch is carboxymethylated by reacting it with sodium chloroacetate in an alkaline medium, followed by neutralization with citric, or some other acid. Cross-linking may be achieved by either physical methods or chemically, by using reagents such as phosphorus oxytrichloride or sodium trimetaphosphate.	Monochloroacetate, nitriles, and nitrates are potentially reactive impurities.	Adsorption of weakly basic drugs and their salts due to electrostatic interactions.[43,44] In addition, the residual monochloroacetate may undergo S_N2 nucleophilic reactions.
Starch	Starch is composed of amylose and amylopectin, polymers of glucose connected by α 1,4 glycosidic linkages (in contrast to cellulose β1,4 linkages). Amylopectin has occasional branch chains connected by α 1,6 glycosidic linkages. Starch is extracted from plant sources through a sequence of processing steps involving coarse milling, repeated water washing, wet sieving, and centrifugal separation. The wet starch obtained from these processes is dried and milled before being used in pharmaceutical formulations. Pregelatinized starch is a starch that has been chemically or mechanically processed to rupture all or part of the starch granules and so render the starch flowable and directly compressible. Partially pregelatinized grades are also commercially available.	Starch may contain formaldehyde, nitrites, and nitrates.	Terminal aldehydes in starch have been known to react with the hydrazine moiety of hydralazine HCl.[45] Starch may also be involved in moisture-mediated reactions, may adsorb drugs, and may react with formaldehyde, resulting in reduced functionality as a disintegrant.[46,47]
Colloidal silicon dioxide		May contain heavy metal impurities.	May act as a Lewis acid under anhydrous conditions and may adsorb drugs.[48,49]
Stearic acid	Stearic acid is made via hydrolysis of fat by continuous exposure to a countercurrent stream of high-temperature water and fat in a high-pressure chamber. The resultant mixture is purified by vacuum-steam distillation and the distillates then separated using selective solvents. Stearic acid also may be made via hydrogenation of cottonseed and other vegetable oil; by the hydrogenation and subsequent saponification of oleic, followed by recrystallization from alcohol; and from edible fats and oils by boiling with NaOH, separating any glycerin and decomposing the resulting soap with sulfuric or hydrochloric acid. The stearic acid is then subsequently separated from any oleic acid by cold expression.		Stearic acid is incompatible with most metal hydroxides and may be incompatible with oxidizing agents. Insoluble stearates are formed with many metals; ointment bases made with stearic acid may show evidence of drying out or lumpiness due to such a reaction when compounded with zinc or calcium salts. A number of differential scanning calorimetry studies have investigated the compatibility of stearic acid with drugs. Although such laboratory studies have suggested incompatibilities (eg, naproxen), they may not necessarily be applicable to formulated products. Stearic acid has been reported to cause pitting in tablets coated using an aqueous film-coating technique; the pitting was found to be a function of the melting point of the stearic acid. Stearic acid could affect the hydrolysis rate of API if the degradation is pH dependent. It could also potentially react with an API containing a primary amine to form a stearoyl derivative.[50,51]

(*Continued*)

TABLE 6.3 (Continued)

Examples of excipients	Method of manufacture	Potentially reactive impurities	Examples of known incompatibilities
Magnesium stearate	Magnesium stearate is prepared either by chemical reaction of aqueous solution of magnesium chloride with sodium stearate, or by the interaction of magnesium oxide, hydroxide, or carbonate with stearic acid at elevated temperatures. The raw materials used in manufacturing of magnesium stearate are refined fatty acids, which is a mixture of palmitic and stearic acid with certain specifications. A fatty acid splitting (hydrolysis) process takes place first, where glycerin and fatty acids are separated. The fatty acids are then further refined to yield tallow acid. Magnesium stearate can be prepared through two processes:1. Fusion—Simple acid-base interaction between tallow acid and magnesium hydroxide2. Saponification—Tallow acid is saponified first with sodium hydroxide, making a sodium tallowate (salt), then magnesium sulfate is added to the sodium tallow solution, followed by pH adjustment, dilution with water, washing, and drying.	Magnesium oxide is a known reactive impurity.	Magnesium stearate can form hydrates with water and exists in four hydration states—monohydrates, dehydrates, and trihydrates.[52] MgO impurity is known to react with ibuprofen.[53] In addition, magnesium stearate provides a basic pH environment and may accelerate hydrolytic degradation.[54] The magnesium metal may also cause chelation-induced degradation.[55]

sulfonic acid.[66] EGTA has a much higher affinity for calcium and magnesium ions than EDTA.

Stabilization of hydrolysis-sensitive drugs intuitively disallows the selection of excipients with high residual moisture content and high water sorption capacity. Nonetheless, excipients with affinity for water might mitigate moisture sensitivity of the formulation by preferentially taking up the moisture permeating through the package during shelf-life and accelerated storage. For example, edible silica gel, Syloid, was used to stabilize extremely moisture sensitive potassium clavulanate in oral solid dosage forms.[67] An interesting case was presented by Perrier and Kesselring of the stabilization of nitrazepam in binary mixtures with excipients as a function of their nitrogen adsorption energy.[15] The authors assumed that the binding energy for water followed the same rank-order correlation as that of nitrogen. The use of excipients with higher binding energy was hypothesized to sequestrate any available moisture and act as a stabilizer. Similarly, it can be hypothesized that the use of excipients that act as sorbents (eg, MCC and amorphous silica) might sequestrate reactive trace impurities (eg, formaldehyde and formic acid and volatile residues such as methanol, ethanol, and isopropyl alcohol) from the formulation.[68] These strategies may be tested in applicable cases by spiking studies in the binary compatibility experiments or the use of mini-formulation designs that combine the offending and the protecting excipients.

Exposure to light can cause drug degradation by various mechanisms such as addition reactions in unsaturated systems, polymerization, isomerization, photo-oxidation, and substitution reactions. The use of light-resistant packaging, for example, amber glass and opaque high-density polyethylene bottles, are standard practices for photolabile drugs. In addition, to protect the drug during processing and packaging operations, light-resistant film coatings and the use of excipients such as cyclodextrins, dyes, and colored additives is often helpful. In addition, photolabile excipients can be used in the formulation of photosensitive drugs just as antioxidants are used for oxidation-sensitive drugs. In this case, substantial overlap of the ultraviolet (UV) absorption spectrum of the excipient with the drug has been shown to improve the drug stability in several cases (eg, stabilization of nifedipine by using riboflavine or curcumin and the stabilization of sulphisomidine by using oxybenzone).[68]

6.3.2 Current practices

Compatibility studies involve a series of activities designed to identify key drug-excipient incompatibilities and their causes (Fig. 6.5). Compatibility studies on new

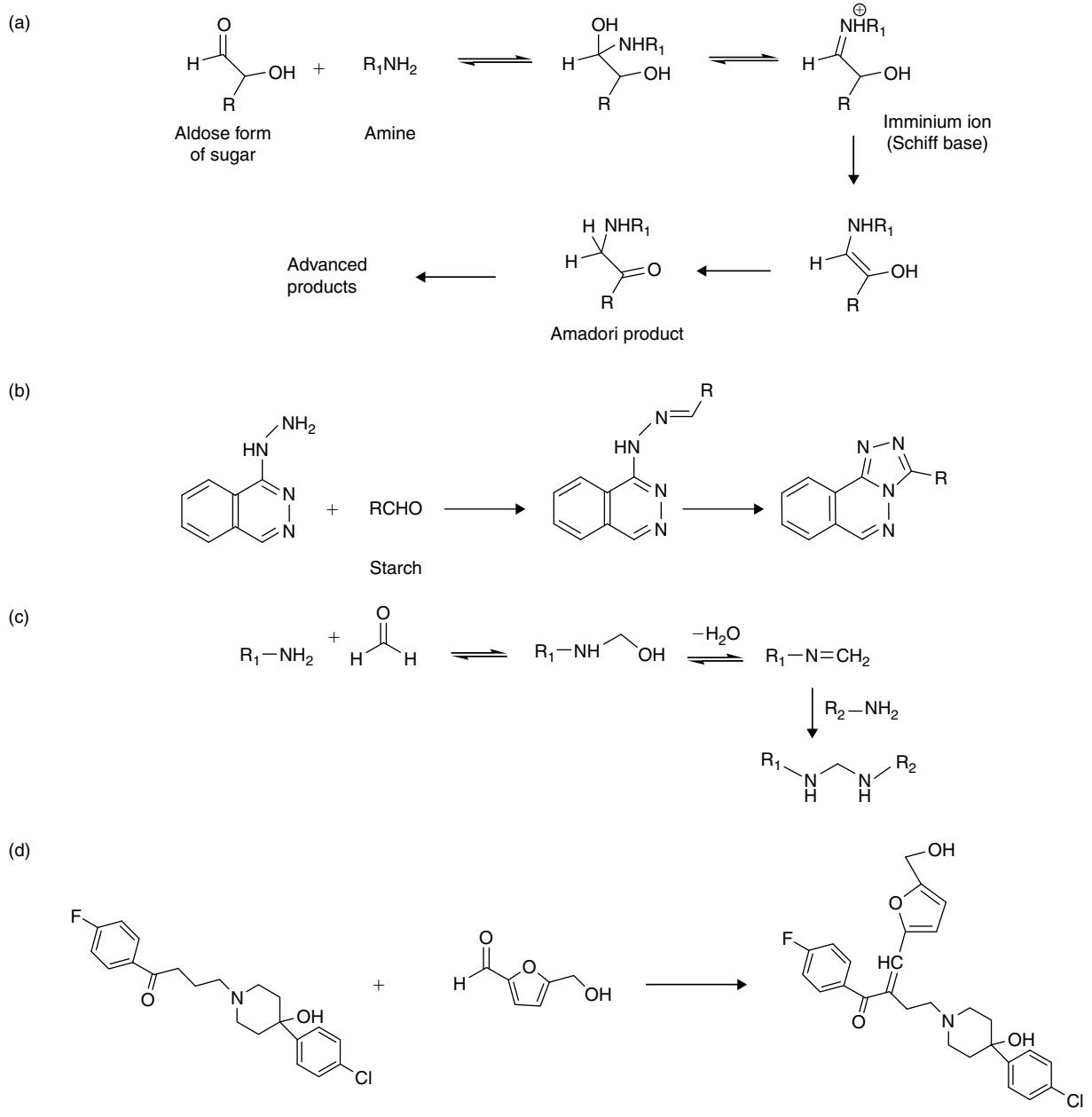


FIGURE 6.2 Examples of reactions of drugs with pharmaceutical excipients and their reactive impurities; (a) reaction of primary amine with a reducing sugar; (b) reaction of hydralazine with terminal aldehydes of starch residue; (c) reaction of an amine with an formaldehyde; (d) reaction of haloperidol with the impurity 5-hydroxymethyl-2-furfuraldehyde (HMF) found in lactose; (e) addition of formic acid to (f) addition of amine to maleic acid; (g) formation of N-oxide from hydrogen peroxide impurity.

molecular entities invariably start with the evaluation of existing information and paper chemistry of the drug candidate to identify "soft spots" in the molecule. The presence of reactive or unstable functional groups, pK_a value, and known reactivity of similar compounds provides useful information for the selection of excipients. In addition to the general literature, several computational programs have recently become available that can help predict potential degradation pathways of a drug candidate (eg, CAMEO, SPARTAN, EPWIN, and Pharm D[3]). Many pharmaceutical companies also have internal databases and software programs. Furthermore, preformulation studies on the physicochemical characterization and forced-degradation of the drug molecule are used to modulate the design of compatibility studies to detect the presence and extent of known reactivity.

The design of compatibility studies might involve the use of mixtures of drugs with one or more excipients. These mixtures may be incubated at different stress conditions as physical mixtures per se or after

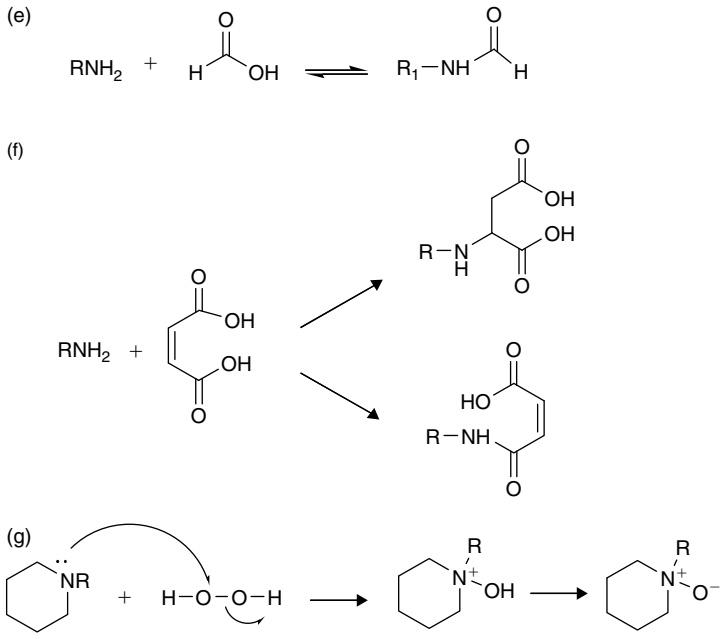


FIGURE 6.2 (Continued)

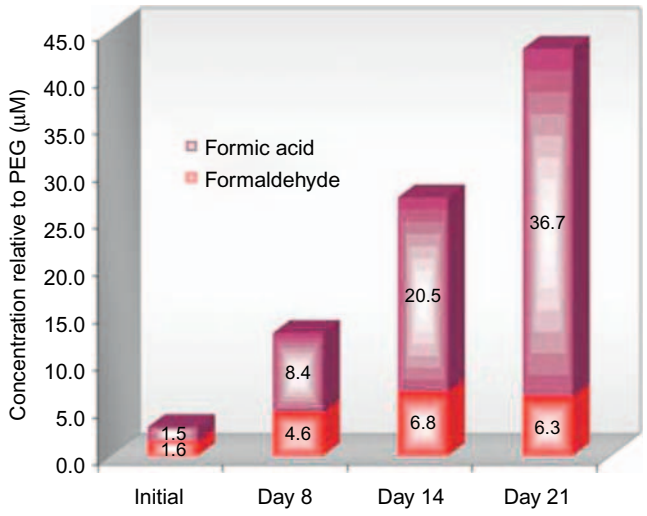


FIGURE 6.3 Generation of formaldehyde and formic acid in poly (ethylene glycol) (PEG) 400 mixed in a 1:1 w/w ratio with water when stored in glass vials with aluminum-crimped rubber septa at 40°C for up to 21 days.[57]

compaction. Often water is added in these systems to evaluate its role in accelerating drug-excipient interactions. The addition of other ingredients, such as hydrogen peroxide to induce oxidative stress, is based on the background information of the molecule's sensitivities. The compatibility study samples are typically stored at elevated temperatures and analyzed for physical and chemical changes in the drug at predetermined time intervals. In addition, binary mixtures of drug and excipients are analyzed by thermal methods such as differential scanning calorimetry (DSC) and isothermal microcalorimetry (IMC) for rapid assessment of potential incompatibilities. In short, compatibility studies involve several choices for each stage of testing depending on the drug candidate, available literature, and the goals of the study. The following sections of this chapter will highlight the basis for some of these decisions.

6.3.2.1 *Experimental design*

Compatibility studies are commonly carried out by accelerated stress testing and evaluation of its effect on the binary or multicomponent drug-excipient mixtures. The design of experiments (DoE) is governed by the potential formulation choices and excipient preferences. These decisions are made in conjunction with all the other available preformulation data, API characteristics, and marketing preferences. These also determine the types of pharmaceutical excipients that are evaluated. For example, compatibility studies for a liquid formulation of an insoluble compound would differ widely (eg, include excipients such as surfactants and suspending agents) from the studies designed for a highly soluble compound.

6.3.2.1.1 Two-component or multicomponent systems

Proactive preformulation compatibility studies are traditionally carried out as binary or ternary systems. Binary mixtures of drug and common pharmaceutical excipients such as diluents, or ternary mixtures of drug, a diluent, and excipients used in lower proportions such as disintegrants and lubricants are incubated at

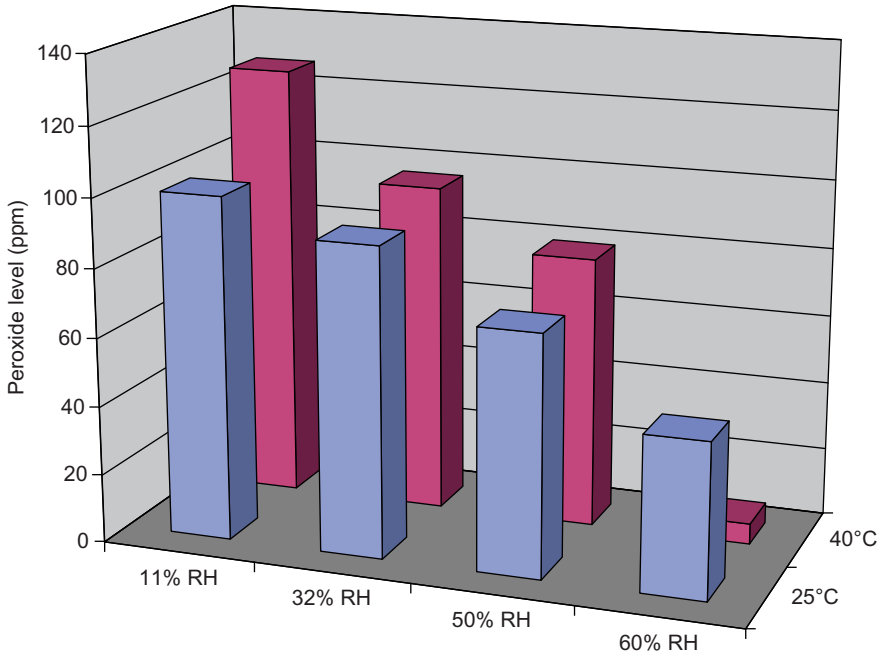


FIGURE 6.4 Effect of storage temperature and humidity (under open-dish conditions) on peroxide levels in povidone (initial level = 80 ppm).[61]

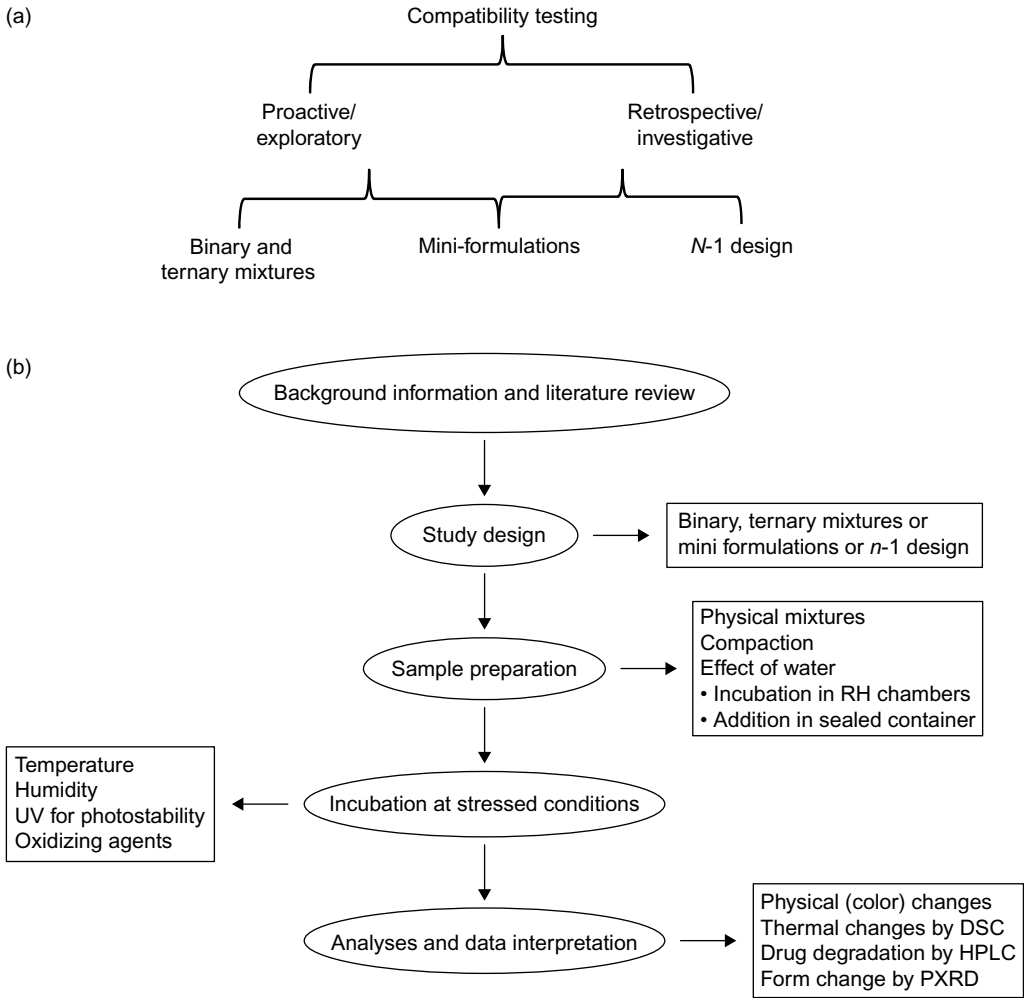


FIGURE 6.5 Typical modalities of compatibility testing (a) and the study execution (b). Various stages of the compatibility testing are highlighted in ovals, and the key decisions and variables involved in each stage are in square boxes.

accelerated conditions of temperature and humidity for extended periods of time, using the drug alone and the excipient alone as controls. Additional aggravating conditions, such as light and peroxides, are incorporated in the study design depending upon the characteristics of the drug molecule. Incompatibilities are physically identified by visual observation for color or physical form changes and spectroscopic and calorimetric methods, and chemically quantified by analytical assays for drug content and impurities.

Wyttenback et al. used binary design to study the excipient compatibility of acetylsalicylic acid or fluoxetine HCl in binary mixtures with 7 excipients using bilevel factors of temperature (40°C and 50°C), humidity (10% and 75% RH), and time (1 and 4 weeks), for a total of 56 experimental runs for each drug.[69] The total impurity content of each run was measured by high-performance liquid chromatography (HPLC) to determine the effect of each excipient on drug stability. In addition, they grouped together all excipients as a factor and interpreted the data by analysis of variance using F-ratio to test whether the means of normally distributed populations are equal, and then calculated the *P*-value. The data presented interesting insights into the relative stability of the two drug substances as a function of these factors and their interactions.

6.3.2.1.2 The $n-1$ design and mini-formulations

Compatibility studies are often aimed at solving formulation stability issues. In such cases, studies are carried out with the exclusion of only one component in each sublot to identify the source of incompatibility. Often, mini-formulations are prepared with the exclusion of noncritical, quantitatively minor, or easily interchangeable ingredients (eg, colors and flavors from solutions and suspensions). Compatibility studies for the development of liquid formulations are invariably mini-formulation studies since they require prior pH solubility and stability evaluation to use the appropriate buffer system in compatibility testing and base formulations.

The Plackett–Burman design may be used to design studies involving mini-formulations, although it is rarely used in practice. This design minimizes the number of experimental runs and is capable of finding the excipients that cause major incompatibilities. It can examine n excipients in $n+1$ experimental runs.[70] This design was utilized by Durig and Fassihi to investigate the compatibility of pyridoxal HCl with 11 excipients at two temperature (25°C and 55°C) and humidity (11% and 75% RH) conditions using only 16 experimental runs.[71] In this study, they included eight experiments over the minimum required to study the effect of "pseudovariables" to account for random experimental variation. This approach, however, does not take into account the variation in the concentration of excipients depending upon the number of components present in the mixture.

6.3.2.2 *Sample preparation and storage*

Sample preparation for compatibility studies depends on the physical nature of the ingredients and the conceptualization of the final formulation. The selection of appropriate drug-to-excipient ratios for binary compatibility studies is often carried out on a weight or molar basis of their expected usage in the final formulation. In the absence of a defined dose of the drug, as is often true of new molecules in the early stages of development, the worst-case scenarios of the lowest drug-to-excipient ratios are tested.

The lowest drug-to-excipient ratio is expected to provide the highest rate of possible drug-excipient interactions. For example, the amorphous calcium salt of the Merck compound L-649,923, a leukotriene D_4 antagonist, degraded by intramolecular esterification to form a γ-lactone. Its degradation rates were higher at lower drug doses with both MCC and pregelatinized starch as excipients.[72] These observations are in line with the hypothesis that reactions happen at the interface of the drug and the excipients. A higher rate of drug degradation with lower-particle-size excipients further supports this hypothesis. Thus, hydrolysis of acetylsalicylic acid in tablets containing dibasic calcium phosphate dihydrate showed a rank-order correlation with the particle size of the excipient.[73] Similarly, powdered MgSt showed higher levels of drug degradation than its granular form.[54]

6.3.2.2.1 Sample preparation

Binary mixture designs for solid-state samples often involve only physical mixing. Attention must be paid to the use of fine particles of both the drug and the excipient, as well as deagglomeration of either component, if needed. Often, coscreening through a mesh is utilized to effect intimate mixing. Compaction of drug-excipient mixtures is carried out in solid-state compatibility testing. Milling and compaction of crystals can lead to the formation of an amorphous state.[74] Process stresses such as compaction, grinding, and drying may also lead to release of bound water from actives and excipients. For example, grinding was shown to dehydrate the crystal water of theophylline hydrate[75] and affected the solid-state stability of ampicillin trihydrate[76] and sodium prasterone sulfate.[77] Similarly, cifixime trihydrate showed a reduction in crystallinity and dehydration upon grinding in a ball mill.[78] Thus, compatibility studies involving grinding and compaction are helpful for moisture-sensitive and low-melting compounds, as is also the case for drug substances for which milling is envisaged for particle-size reduction.

The presence of water often accelerates degradation kinetics. In addition, sorbed moisture can accelerate isomerization, crystallization, or both. One of the important goals of preformulation compatibility testing is to determine the feasibility of processing conditions such as wet granulation and the moisture resistance need for packages. Thus, it becomes important to determine the case and extent of moisture sensitivity of the drug in the presence of excipients. To this end, water is often added to drug-excipient binary mixtures in closed systems to create a high-RH environment.[78] Alternatively, the compatibility mixtures are stored in open containers at different temperature and RH conditions. Care must be exercised in comparing the data across these designs since the potential volatilization and escape of reactive components in the open-container storage conditions might give an indication of stability that may not hold true in the final dosage form.

Compatibility study experiments tend to be labor intensive due to their exhaustive design. Automation of compatibility studies is often undertaken during the early development stages.[79] The automated systems can be used for determining drug solubility, forced degradation, and compatibility testing. Reactor blocks, consisting of an array of samples, can be prepared using automated liquid handlers and powder dispensers such as Autodose and Gilson liquid handler. The blocks can be stored in different stress conditions. Automated sampling at predetermined time points can be accomplished using robotic systems such as Jaba. These operations also may be performed with fully automated systems such as the Symyx Automated Forced Degradation System. This system can accurately weigh given quantities of excipients and actives in blocks of vials. The blocks are then robotically transferred to controlled environmental chambers and removed at a predetermined time for analysis. The use of automated systems improves efficiency and broadens the experimental space, allowing the execution of exhaustive, statistically robust experiments and generation of knowledge. This, when coupled with scientific rigor, could significantly reduce or even avoid the appearance of drug-excipient incompatibilities during later stages of DP development.

The selection of accelerated environmental storage conditions for compatibility testing is based upon the nature of the API and the expected stresses during the storage of the finished dosage form. These include storage at elevated temperatures, higher humidity, and exposure to UV irradiation, peroxides, etc. While the International Council on Harmonization (ICH) and the regulatory authorities provide guidelines on the selection of accelerated storage conditions for the finished dosage form for confirmatory stability studies based on the target label storage requirements, the selection of accelerated conditions for compatibility studies is at the formulator's discretion.

6.3.2.2.2 Thermal stresses

Isothermal stress testing (IST), or incubation of samples at constant higher temperatures as stress conditions, is almost universal for compatibility studies. It is based on the underlying assumption that the kinetics of degradation reactions follows Arrhenius kinetics. The Arrhenius equation states the reaction rate dependence on temperature:

$$k = A \times e^{-E_a/RT}$$

where T is the temperature (in Kelvin), R is the gas constant with a value of 8.314 J/(K × mol), E_a is the activation energy in KJ/mol, k is the reaction rate constant, and A is the Arrhenius constant.

The selection of higher temperatures is designed to accelerate reaction rates significantly so that even relatively slow reactions become evident in a short period of time. For example, assuming the mean activation energy for solid state reactions to be 105 kJ/mol,[80] the amount of drug degradation observed at 25°C in approximately 5 years would be evident in about 3 weeks at 60°C. Certain implicit assumptions of IST include:

- Temperature independence of activation energy and reaction mechanism
- Absence of equilibrium or autocatalytic reactions
- Activation energies for different reactions of pharmaceutical relevance in a relatively close range
- Stability of API per se under the stress conditions.

The temperature is often maximized to increase reaction rate and decrease testing time. However, care must be exercised in the judicious selection of temperature since above a certain temperature, the system may exceed the activation energies of alternative degradation pathways, resulting in nonrepresentative data. A shift in the primary degradation processes is often indicated by the appearance of additional impurity peaks in the HPLC chromatogram. For example, while exploring accelerated degradation conditions for an experimental compound, Sims et al. observed that while temperatures up to 80°C resulted in the formation of the same impurities as controlled room temperature and 60°C, samples stressed at 100°C showed many unrelated peaks.[81]

Therefore, usually more than one stress condition is chosen to determine if the degradation reactions observed at elevated temperatures might be the ones that have such high activation energies that they occur only under pharmaceutically irrelevant conditions.

6.3.2.2.3 Humidity and water content

The use of high moisture content for compatibility studies is intended to uncover not only the reactions in which water participates directly (eg, hydrolysis), but also to investigate whether water increases reactivity in the solid state. Water adsorption on solid surfaces

can enhance reactivity by acting as a medium or a plasticizer for the reacting species. The presence of water in the solid component mixtures provides the molecular mobility required for reactions to occur between the two solid components.

Water can be incorporated in the excipient compatibility samples in several ways: (1) preparation of a slurry or suspension; (2) addition of water, usually 20% of solids, in a closed system[81]; and (3) exposure of the system to controlled humidity conditions.

Each of these methods has advantages and disadvantages. While the slurry experiments provide useful information for the simulation of suspension dosage forms and rapid assessment of aqueous sensitivity of the drug substance, they often do not simulate conditions in solid dosage forms, which could be important to data interpretation in retrospective or investigational compatibility studies. The addition of a fixed proportion of water to the drug-excipient mixture in a closed system controls the initial amount of water in the system and partly simulates wet-granulation conditions. Nevertheless, the water content in a sample changes depending on the water uptake kinetics and equilibrium moisture content of the sample at the storage temperature. Moreover, the same amount of water in the air results in lower RH at elevated temperatures. In contrast, incubation of samples under controlled humidity conditions ensures consistent maintenance of equilibrium moisture levels in the solid system, enabling the investigation of degradation reactions where water acts as a medium or just to increase the molecular mobility of reacting species.[81] However, the open-container studies may lead to loss of reactive volatile impurities from the mixture, possibly altering the reaction pathways. In addition, storage of samples in constant humidity chambers (eg, open dish) can lead to increased induction times compared to water addition (or slurry experiments), thereby complicating kinetics.

Given the pros and cons of each modality of testing, the formulation scientist must select appropriate conditions for each compatibility study design based upon the available preformulation data.

6.3.2.2.4 Mechanical stress

Mechanical stress is often an unavoidable part of DP processing conditions (eg, milling and compaction). These stresses may lead to the formation of amorphous pockets of drug or crystal defects in the crystalline drug molecule, or simply increase the intimacy and surface area of contact between the drug and the excipients in the mixture. Thus, mechanical stresses may lead to increased rates of degradation reactions. For example, milling of Procaine Penicillin G almost doubled its rate of degradation in compatibility studies with Emcompress and Avicel as compared with the unmilled drug-excipient mixtures.[82] Similarly, Badawy et al. observed more than threefold higher hydrolytic degradation of an investigational compound DMP-754 in compacts with anhydrous lactose as compared to uncompacted binary mixtures.[83]

In addition to chemical reactivity, compaction may have an effect on the physical form of the drug. Thus, Guo et al. observed a loss of crystallinity of quinapril hydrochloride, whose crystalline form exits as an acetonitrile (ACN) solvate, with ACN loosely incorporated into the crystal lattice channels. Loss of ACN was observed upon compaction, with concomitant loss of crystallinity and increase of the cyclized degradant. This was attributed to the significantly reduced activation energy for the cyclization reaction in the amorphous state than the crystalline state.[84]

Investigations of the effect of mechanical stresses are undertaken when instability is suspected in the investigational drug (eg, in situations such as the use of a metastable polymorphic form or known chemical reactivity of the molecule). Simulation of mechanical stresses in compatibility studies involves such procedures as grinding using a ball mill and comilling of drug-excipient mixtures. In addition, a Carver press for the preparation of compacts is often utilized to simulate the compaction during tableting and dry granulation.

6.3.2.2.5 Oxidative stress

Oxidation is one of the most common causes of drug degradation in pharmaceutical systems. Oxidative degradation often presents unique patterns of drug degradation (eg, rapid growth of impurities following an induction period and sensitivity to trace amounts of free radicals). If preformulation or API characterization studies reveal sensitivity to oxidation, the compatibility studies may be designed to comprehensively assess the extent of oxidation sensitivity as well as to find ways to mitigate the same. These methodologies include, for example, one or more of the following:

- Spiking drug-excipient mixtures with different levels of oxidizing agents, such as hydrogen peroxide; metal impurities (copper and iron salts); or free radical initiators commensurate with the levels seen in excipients[85]
- Comparing drug degradation in compatibility samples stored under air, oxygen, and nitrogen or argon in hermetically sealed containers
- Incorporation of free-radical scavengers and heavy metal chelating agents[85]
- Packaging configurations that include the incorporation of oxygen scavenger inserts and the minimum permeability materials of construction[86]
- Use of different batches of excipients known to have residual peroxide content (eg, PVP, cross-linked PVP, HPC, polysorbate 80, and PEG 400)[38]

- Studies with different antioxidants (eg, BHA, BHT, α-tocopherol, propyl gallate, and ascorbic acid)[87]

For these studies, sample preparation conditions, such as compaction, should be carefully controlled. Furthermore, peroxide impurities in excipients usually show batch-to-batch variation, thus necessitating the testing of batches with high- and low-level impurities. Spiking experiments are often carried out to quantify the acceptable limits of peroxide impurities in excipients, which may then be used to set incoming material specifications.[38]

6.3.2.3 Sample analysis and data interpretation

The desired tests for a given set of compatibility studies and the key outcome measures depend not only on the envisaged final dosage form and product configuration, but also on the background data available on the chemistry and preformulation studies on the drug candidate. Most compatibility studies include visual inspection for any color changes, compact/tablet integrity, and deliquescence; and quantitative chemical analysis for monitoring drug degradation. In addition, form changes of the active are monitored in samples of short-listed excipients.

6.3.2.3.1 Monitoring for drug degradation

Physical observation of stressed compatibility samples involves the observation of changes in color, odor, deliquescence, powder/compact flow characteristics, and other characteristics. The inherent subjectivity in these observations is partly overcome by making rank-order correlations with control, refrigerated, samples. Furthermore, UV-visible spectroscopy may be utilized to quantify changes in color. This method, however, suffers from the limitation that small changes in absorbance may not appear to be significantly different and the degradants may retain the chromophores of the drug, leading to negligible or no change in absorbance even in the presence of degradation. Nevertheless, when significant differences are seen, these observations are often a good indicator of incompatibility between the components.

Several drugs show such evidence of discoloration as an indication of instability, incompatibility, or both, including promethazine, phenylephrine, potassium clavulanate, cefuroxime axetil, and terbinafine. A common example is the discoloration (often browning) of formulations containing primary or secondary amines with reducing sugars such as lactose.[88] This is associated with the Maillard reaction, whose end products can undergo Amadori rearrangement to form several colored intermediates and end products.[56]

HPLC with UV detection is by far the most commonly used method to quantify drug degradation by measuring drug potency, total impurities, or the growth of a selected impurity over time and as a function of the storage conditions. Measurements of growth of a single impurity or all impurities are often preferred over the quantification of loss in potency. This is because a small change in a large number reduces the sensitivity in the interpretation of the data; for example, a potency drop from 99.5% to 99.2% would likely be considered insignificant, but not an impurity level increase from 0.02% to 0.32%.

It is important that the sample extraction, along with the analytical methodology, is sound. Particular attention should also be paid to the lack of mass balance issues, which can arise either due to poor extraction or retention of degradation products on the column or due to reduction in the response factor of the degradation products. When extraction problems are suspected, alternate solvents and agitation techniques should be utilized. To ensure that the column retention is not an issue, HPLC systems with a wide polarity range and longer run times can be employed. If the mass balance issue is severe, the relative response factors of drug and impurities may be assessed by taking the ratio of absorbance of the degraded sample to that of the initial (reference) sample measured by UV alone. This ratio should be identical to the ratio of total area counts for the degraded and initial sample as determined by HPLC-UV. The main disadvantage of this method is that it is useful only when there is a severe mass balance issue. Although not used in routine analyses, alternative detectors such as evaporative light scattering detector, refractive index detector, mass spectrometry, Corona charged aerosol detector, chemiluminescent nitrogen detector, or LC/NMR can be used when UV detection is not effective.

6.3.2.3.2 Thermal methods

The use of a quantitative, stability-indicating analytical method such as HPLC may not be feasible during the early stages of drug development. In such cases, thermal methods may provide a quick, nonspecific, and less labor-intensive screening tool. It must be noted that results obtained from thermal methods are often not conclusive and must be confirmed by another independent observation.

Thermal methods for compatibility testing rely on exothermic or endothermic energy changes in the sample. The most commonly applied thermal technique for compatibility studies is DSC. It involves heating or cooling the sample in a controlled manner and measuring the heat released or absorbed by the sample as the temperature of the sample and a reference standard are equally changed over time. IMC, on the other hand, measures heat flow from the sample compared to a reference vessel as the two are maintained at a

constant temperature. IMC can be applied to detect extremely slow reactions since this technique allows the measurement of very low amount of heat changes. This allows the ability to monitor reactions in drug-excipient mixtures at room or low temperatures. Thus, reactions that are relatively slow and affect the formation of impurities in DP over prolonged shelf-life storage (ie, reactions of real-life significance) can be monitored using IMC.[89,90]

DSC analyses for compatibility evaluation involves recording the thermograms of the individual excipient, the drug, and their physical mixture at a standard heating rate, usually under a nitrogen atmosphere. A simple superimposition of these curves is then interpreted to determine whether the thermal properties of the mixture are a sum of the individual components, an assumed fundamental property of noninteracting components. An interaction is, thus, identified as changes in the appearance or disappearance of a transition peak and through changes in transition temperature, peak shape, and peak area. Nonetheless, caution should be exercised in interpreting the results of DSC for determining excipient compatibility and it should be in conjunction with simultaneous application of other techniques such as infrared (IR) spectroscopy and IST. For example, Verma and Garg utilized physical observations, IR spectroscopy, and IST in conjunction with DSC to determine excipient compatibility.[91]

IMC can measure thermodynamic events in powder mixtures at real-time storage conditions, thus potentially pointing to the degradation reactions that are of relevance in the long-term stability studies of the DP. For example, it can pick up reactions that may lead to only 1−2% degradant accumulation over a year of storage at room temperature. The high sensitivity of this technique necessitates careful control of sample preparation (eg, the uniformity of particle size of different samples, adequate mixing, and the effects of moisture). Schmitt et al. used IMC to study the compatibility of an experimental compound ABT-627 in binary mixtures with excipients.[92] Selzer et al. reported IMC studies on drug-excipient mixtures as powders, granules, and compacts.[93,94]

6.3.2.3.3 Monitoring for form changes

The drug discovery and early-stage development work is aimed at selecting not only the right molecule, but also its most preferred solid form. The physical form of the drug, polymorphic and salt form, has a significant impact on the physical, chemical, mechanical, and biopharmaceutical properties of the drug candidate. For example, a state with higher thermodynamic activity (eg, the amorphous phase) generally has higher apparent solubility and diffusion-controlled dissolution rate, as well as higher hygroscopicity and chemical reactivity. It also has a tendency to crystallize to a more stable form. The mechanisms underlying phase transformations have been discussed in the literature.[95,96] It is important to ensure that dosage form composition and processing conditions do not alter or destabilize the physicochemical form of the drug. It must be noted that although these investigations are not a routine part of compatibility testing, they become very important if a metastable form is taken up for development or if there is data on the API stability that warrants monitoring for form changes.

Several analytical techniques can be utilized for the characterization of drug's form in the solid state, for example, powder X-ray diffraction (PXRD), single crystal X-ray diffraction, IR and near-infrared (NIR) spectroscopy, Raman spectroscopy, solid-state NMR spectroscopy, solvent sorption, polarized microscopy, and hot stage microscopy. An example of the application of instrumental techniques to identify the impact of excipients on potential form change of the drug during wet processing was presented by Airaksinen et al. using nitrofurantoin as a model drug.[97] Nitrofurantoin exists in two anhydrous (designated α and β) and two monohydrous (designated I and II) forms. High-humidity storage and processing conditions (eg, wet granulation) may lead to the conversion of the anhydrous to the monohydrate form. The authors used PXRD and NIR spectroscopy to investigate form transformation of nitrofurantoin in physical mixtures with selected excipients. Their interpretation of PXRD data focused on the appearance of a peak specific to the monohydrate form that indicated form conversion. Similarly, NIR data was interpreted for the appearance of the absorption maxima for the water of crystallization in the monohydrate form. Such data sets help in establishing the rank order relationships of the tendency of excipients to support or inhibit form conversion of the drug. The authors correlated the water sorption capacity and the amorphous/crystalline nature of the excipients with their role in moisture-induced form change of nitrofurantoin during processing. Their work demonstrates the utility of these techniques in investigating drug form changes as a part of the compatibility studies.

6.4 EXCIPIENT VARIABILITY

The drug-excipient and excipient−excipient interactions in dosage forms can significantly affect both the stability[98] and the bioavailability[99] of the drug substance or the API.[100] The mechanistic basis of such effects could be a direct interaction between the drug and the excipient, reaction of an API with the reactive impurities that may be present in excipients,[101,102] or indirect effect of excipients on DP attributes such as the water

activity of the dosage form[103] and its densification.[104,105] Variability in excipient properties that contribute to their effect on DP attributes can, therefore, lead to batch-to-batch variability in DP QAs. To obviate these possibilities, as a rule of thumb, simple formulations that minimize the use of excipients to essential functional ingredients is preferred.[106] In addition to a thorough mechanistic understanding of physicochemical interactions in a DP, identifying relevant excipient properties that contribute to critical product attributes and delineating their normal acceptable and generally observed variation is critical to robust DP design.

QAs of excipients do vary—not only between different grades and types of excipients, but also within the same excipient. Variability of excipient attributes could be source-to-source variability or lot-to-lot variability for the same grade and source of excipients. These have also been termed *intersource* and *interlot variability*, respectively. Intersource variability is much easier to control. It is usually done by establishing and controlling the source of raw materials. Interlot variability, however, can be difficult to identify, understand, and control, especially when these are outside the types of material properties reported on the vendor's certificate of analysis (CoA). Such properties may include, for example, compression properties of Elcema G250, an α-cellulose product,[107] lubrication properties of MgSt,[108–110] and compaction properties of starch.[111] The extent of lot-to-lot variability could be different from one supplier to another. For example, when Williams et al.[112] studied different particle-size grades of MCC using six different suppliers and two lots per supplier, they observed greater lot-to-lot differences with the material supplied from some suppliers than others. This section primarily focuses on the interlot variability of excipients and suggests a set of approaches that can be proactively adopted to understand and account for the lot-to-lot variability encountered with excipients when designing robust DPs and input material control strategies.

Addressing excipient variability during DP development has the following components[113]:

1. Identification of critical excipients in a given DP
2. Understanding of the mechanistic basis of functional role of the critical excipients
3. Identification of CMAs of the excipient that affect DP performance
4. Developing an understanding of the range of excipient QAs that are or may be routinely encountered during DP development
5. Building a DP design space or QbD strategy that incorporates the observable differences in excipient QAs

6.4.1 Identification of critical excipients

Pharmaceutical development of new DPs routinely address the effect of formulation composition (both qualitative and quantitative) and process parameters on DP QAs.[6] Excipients are utilized in the DPs to aid in the bioavailability, stability, or manufacturability of a dosage form. Most functional excipients utilized in the DP fulfill one or more of these roles. At the same time, the criticality of an excipient is usually assessed with respect to the influence of variability in excipient attributes on the patient-centric QAs. Certain excipients, such as colloidal silicon dioxide, are added to oral solid dosage forms as functional excipients that help with the flow and prevention of sticking of the granulation during manufacture. These excipients are important for developing a robust DP manufacturing process. At the same time, they may not directly affect a patient-centric QA of the DP. Therefore, it is important to determine the excipients that are critical with respect to patient-centric QAs and differentiate these from the excipients that may be critical from the process robustness or manufacturability perspectives. Early identification of excipients critical to the DP manufacturability (such as process robustness) can influence the selection of formulation components to not only build robust processes, but also reduce the impact of excipient variability on formulation QAs or manufacturability.

6.4.2 Understanding the mechanistic basis of functional role

Functional excipients are typically added to the DP for a well-defined objective. For example, immediate release tablet dosage forms typically include a disintegrant to promote rapid tablet and granule disintegration to facilitate drug release. Sustained release tablets may have a hydrophobic core matrix to slow the rate of drug release. Formulation and process development studies are typically carried out to determine the effect of excluding or replacing a particular excipient (qualitative composition change) or the concentration of an excipient (quantitative formulation change) in controlled, often statistical, DoE studies. These studies result in a robust definition of factor-response relationships with an additional delineation of two-way or three-way interactions. Such studies, however, do not provide adequate information about the critical attributes of the excipient that is important to its functionality.

The mechanistic basis of the functional role can be better understood using excipients in the same

functional class, but with well-controlled and -defined differences in material properties. For example, certain excipients can show noncovalent interactions with drugs in the dosage form that can affect drug release in the dissolution medium. The effect of drug-excipient binding on a drug's bioavailability, however, would depend on the strength and extent of binding of the drug. For example, CCS, a common superdisintegrant in immediate-release oral solid dosage forms and an anionic polyelectrolyte excipient, can form noncovalent interactions with basic amine drugs. A source of interlot and intervendor variability in CCS stems from the degree of cross-linking within the polymer network. A comprehensive set of modeling and experimental studies[8] indicated that the relevance of drug-excipient binding to a drug's bioavailability depends on the drug's dose and the affinity of binding (Fig. 6.1b).

6.4.3 Understanding the range of variability of excipient attributes

A DP manufacturer has very little control over the intrinsic lot-to-lot variability within a single grade of each excipient in a given formulation, as the source of this variability is dictated by the degree of process control implemented by each excipient vendor.[114] Several lots of excipients are typically used during formulation design, formulation and process development, process scaleup, and clinical DP manufacture, which sample the range of typical variability expected in the excipient attributes. In addition, several lots of commonly used excipients are routinely procured and utilized by the pharmaceutical industry. The exposure to different lots and manufacturers of excipients provides opportunities to test the range of variability observed in different lots of excipients for a particular product QA. The quantitation of the range of excipients' QAs encountered during the development of a DP may be used to build a baseline database of excipient attribute experience-space[115] that can be helpful for establishing formulation and process robustness, and to study the effect of potential deviations from this experience space in the future batches. Risk assessment may be used to shortlist the attributes to be monitored in building the database of excipient attribute experience-space.

A common approach is collating data, and establishing statistical parameters, from the information provided in the CoA by the vendor. This approach can sometimes be combined with Bayesian statistical methods, which can enable continuous development of a statistical database as representative batches of the DP are manufactured over a period of time. For example, Kushner et al. reported discernible variability of Avicel PH 102 manufactured at two different sites although all data was within specifications.[116]

Developing an understanding of the effect of excipient attributes on DP MAs represents an inherently multivariate to potentially more than one response. The use of chemometric methods with statistical multivariate analyses tools such as principal component analysis and multiple linear regression have been proposed for this purpose.[117] These tools can allow better understanding of the effect of MAs on in-process material properties, which are ultimately linked to DP CQAs, and process outcomes—thus enabling the development of a robust multivariate advanced process control strategy.

6.4.4 Generating or obtaining excipient lots with a range of known MAs

A traditional approach to determine the impact of excipient QAs on DP performance is to select the known MAs that may affect product performance based on literature, functional role of the excipient in the formulation, or prior experience, and to investigate the effect of a range of those excipient attributes on product CQAs in controlled experimental studies. Such an experimental investigation of the effect of selected *potential* CMAs of excipients on predefined DP CQAs relies on the ability to obtain excipients with a wide-enough range of those potential CMAs. For example, to examine the impact of excipient material property variability for three common excipients on the performance and manufacturability of an immediate-release, dry granulated, solid oral dosage form, Kushner et al. selected the material QAs that were known to affect product performance.[114] These were the particle size of MCC, spray-dried lactose, and MgSt, as well as the polymorph and specific surface area of MgSt. The authors obtained samples of different batches of these excipients from vendors with differences in selected QAs for this study. The authors observed that the particle size of MCC and lactose affected blend particle size, blend flow, granulation particle sizes, roller compaction ribbon tensile strength, and tablet hardness. Specific surface area of MgSt affected ribbon tensile strength and tablet hardness. The authors concluded that despite the observation of statistically significant effects, the range of particle sizes examined were not sufficiently large to adversely affect the CQAs studied.[114] This conclusion highlighted that even though variation in an MA had a statistically significant impact on in-process MAs during DP manufacture, it did not adversely affect the CQAs of the finished DP.

A frequently cited limitation of this approach is the availability of a given excipient with sufficiently large variation in the QAs that can cover the range of variability listed as a specification on the official CoA provided by the vendor. The manufacture of excipients in bulk using highly controlled processes results in a few batches of excipients manufactured annually and those batches being very tightly controlled in terms of the QAs under investigation. Two approaches are suggested when such limitations make experimental verification of excipient variability impact on DP CQAs challenging:

- Using different grades or suppliers of excipients that would represent greater variation in the selected MAs than within grade, thus qualifying as a worst-case scenario for the selected MA
- Spiking or using storage conditions that result in modifications to the MAs under investigation

6.4.4.1 Different grades or suppliers of excipients as a worst-case scenario

When it is difficult to obtain samples of different batches of an excipient with adequately wide differences in selected MAs *and* such MAs vary across different grades of an excipient, different grades of the excipient can be selected for an experimental investigation of the role of selected MAs on potential DP CQAs and manufacturability as a worst-case scenario.

A caveat of this approach is that the range of variation in the MA examined may be wide enough to significantly affect DP CQA or manufacturability when the range of variability routinely expected within a grade of an excipient may not have affected DP CQA or manufacturability. For example, Alvarez-Lorenzo et al. described interlot (ie, two lots) and intersource (ie, two suppliers) variability in the properties of HPC and its implications on drug release from theophylline swellable matrix tablets.[118] The authors compared two lots of a high-viscosity grade of HPC, which showed differences in the molecular weight, molecular structure, particle size distribution, particle shape, and water affinity. These parameters were seen to affect the rate of drug release from tablets, with the impact of interlot variability being lower than intersource variability. In another study, Doelker et al. investigated the tableting characteristics of 16 different lots of MCCs produced by 7 different manufacturers.[119] These grades were compared with respect to moisture content, particle size distribution, bulk and tap densities, flow properties, and sensitivity to lubricant. The authors found greater intersource than interlot differences. These studies indicate that the results of studies conducted with different grades or suppliers of excipients as a surrogate for the extent of lot-to-lot variability within a grade or supplier should be interpreted with caution. This approach (intersource or intergrade variability as a worst-case scenario for interlot variability) is generally more valuable for ruling out a potential MA as a CMA than identifying a particular MA as a CMA.

6.4.4.2 Mixtures of different grades of excipients

A modification of the approach of using different grades of excipients to delineate the impact of wide variations in MAs on product QAs is to use the mixtures of excipients that provide material properties within the range of the specifications of target material. This approach was utilized by L'Hore-Gaston et al.,[120] who prepared physical mixtures of different grades of polyethylene oxide excipients to simulate product viscosities across the range of the standard product specification of the target material the authors were planning to use. The authors used two standard poly (ethylene oxide) polymers (POLYOX 205 NF Water-Soluble Resin and POLYOX N-12K NF Water-Soluble Resin) to develop a series of samples across the viscosity specification range of another standard product (POLYOX 1105 NF Water-Soluble Resin), which has a viscosity between these two grades.

For this study, the authors had preidentified the MAs to study (polymer viscosity and particle size distribution), the DP of interest (an extended release matrix system), and the product QAs as responses (tablet mechanical strength and drug dissolution).[120] The approach of mixing the two different grades of excipients as producing results similar to an intermediate viscosity grade was verified by comparing drug dissolution profiles from matrix tablets. Utilizing this approach, the authors were able to make reasonable inferences of the properties of the polymer on drug release from the matrix extended release tablet. Where feasible, such an approach can be a valuable tool to study the acceptable range of variation in material properties that can result in a truly QbD raw material control strategy.

6.4.4.3 Spiking or using storage conditions to modify MAs

Reactive impurities in excipients that react with the drug substance or API in a formulation can lead to safety concerns about the drug degradants. This section describes potential methodologies to test the effect of variability in excipient impurities on DP stability when excipients with different concentrations of reactive impurities are not readily available from the vendor. These approaches consist of storing excipients under particular conditions or spiking excipients with known reactive impurities. For example, reactive

formyl impurities in PEG increase upon storage at high temperature (Fig. 6.2).[57,121] Reactive peroxide impurities in povidone increase upon storage under low-RH and high-temperature conditions.[61] Thus, storage of excipients under conditions that lead to an increase in the reactive impurities under investigation can be useful to experimentally study the effect of reactive impurities in excipients on DP stability in a controlled manner, using the same lot of excipient stored under different conditions.

The strategy of spiking known reactive impurities in excipients has been utilized in certain cases such as spiking hydrogen peroxide in povidone and crospovidone or formic acid and heavy metals in PEG, polyvinyl alcohol (PVA), or coating materials comprised of these polymers (eg, Opadry II from Colorcon). Spiking strategies, however, have inherent limitations. For example, the reactive impurities' species present in the excipient resulting from natural formation under routine manufacture and storage conditions can be very different and diverse compared to the spiked species. For example, spiking of formic acid in PEG or Opadry II coating material may not reproduce the variety and extent of formyl esters present in the polymer networks when the impurities are formed inherently. Spiking of formaldehyde has practical limitations, in that formaldehyde is a gas available as a 37% aqueous solution. Spiking of hydrogen peroxide in povidone does not address the organoperoxide impurities that are present inherently. In addition, the uniformity of distribution of minor quantities of spiked impurities within the excipient sample and how well or deeply the spiked impurities are embedded within the polymer network is always questionable.

6.4.5 Controlled experiments with a range of known MAs

Once different lots of excipients have been obtained with differences in MAs of interest, experimentation is undertaken to investigate the impact of the known differences in MAs on the selected product QAs.

6.4.5.1 MA comparison at target formulation and process parameters

Investigation of the effect of MAs—of one or more components of a formulation—is typically carried out without studying the interaction with the formulation composition or process parameters. In most cases, the effect of changes in the MAs are studied at the center point of the target formulation composition and process parameters. These experiments seek to compare the responses due to potential excipient lot-to-lot variability, while keeping all other formulation and process parameters constant. A limitation of such experiments is lack of information about the implications of interactions of variability in more than one excipient or the interaction of excipient variability with process parameters. These studies can sometimes incorporate the variability in more than one excipient to study their interactions. In certain cases, focused experiments are carried out to study the impact of two or three variables simultaneously. For example, the effect of surface area or particle size of MgSt in two different lots can be studied concurrently with the number of revolutions during the lubrication mixing operation in a 2^2 full-factorial design to identify the interaction between the formulation (lot-to-lot variability) factor and the process factor.

6.4.5.2 Statistical DoE studies that combine MAs with formulation and/or process parameters

Generating an extended experimental space using systematic statistical DoE with different lots of several excipients combined with some formulation and or process variables, although feasible, becomes prohibitive in material, time, and effort requirements. The feasibility and challenges encountered in such a study were highlighted by the work of Gabrielsson et al.,[122] who used multivariate design to study the robustness of a tablet formulation. The authors investigated the impact of lot-to-lot variability in excipients in a formulation consisting of 10 ingredients (9 excipients and 1 API). These studies were conducted at two different drug loads. Utilizing 3–7 lots of each of the 10 ingredients, the authors manufactured and characterized 50 DP batches in a first multivariate characterization study. Another multivariate characterization was performed on 30 batches, 3 from each of the 10 components of the formulation. The results of this study were subject to principal component analyses to determine the principal properties. A Plackett–Burman design was applied to the principal properties. Regression analysis was used to study the relationship between design factors and two selected DP responses (mean disintegration time and mean crushing strength). The relatively lean saturated statistical study design allowed the delineation of only the main effects. These studies allowed the authors to generate multivariate characterization models for each of the main effects studied in the design.

Another example of multivariate statistical DoE study that incorporated excipient attributes with formulation and process variables is provided by Kushner et al.,[123] who studied within grade and intergrade variations in excipient particle size along with API type, API particle size, drug loading, manufacturing method, and a range of diluent to lubricant particle size ratios in a 2^{5-1}

fractional factorial statistical design. The authors conducted 28 experimental runs with a variety of response measurements, and noted that the variation in excipient particle size distribution parameter, reported d[4,3] value, had the greatest impact on the blend and granulation particle size and flow responses.

DoE studies that encompass formulation and process variables along with the excipient attribute differences can be relatively large and demanding in terms of material, time, and effort requirements. At the same time, these studies also tend to be limited in terms of the variables that can be investigated. The investigator needs to select the most pertinent formulation and process variables for study, as well as the excipients within the formulation and their range of variables that can be studied.

To design a robust DP formulation, pharmaceutical scientists need to take into account the physicochemical properties and variability of the API; the performance advantages, limitations, and variability of the excipients; the advantages, limitations, and variability of the unit operations; the influence of input and in-process material storage conditions and DP packaging components; and interactions among these variables.[124] Thus, a holistic understanding of influence of excipient impurities on DP QAs can allow the development of a control strategy with the DP focus. For example, for excipients that show increase in the concentration of impurities during storage, such as PEG[57,121] and povidone,[61] the packaging configuration and storage conditions may be more fruitful than tight specifications on reactive impurity concentration in the initial input material. An increase in reactive impurities in excipients during storage stability in a DP can sometimes be mitigated by avoiding the concurrent use of excipients that can increase the rate of formation of certain reactive impurities through excipient–excipient interactions, as, for example, the use of PVA and PEG in conjunction in a drug product.[125,126]

Process parameters may interact with excipient attributes to generate reactive impurities in DP that can be controlled by establishing adequate process control strategies. For example, Reed et al. reported on increase in organoperoxide-mediated degradation in a DP that correlated with fluid bed granulation process parameters[127] and the photosensitivity of a drug that correlated with light exposure during manufacturing and packaging processes in addition to the concentration of iron in the excipients.[128] Hemenway et al. demonstrated that, in certain cases, the rate of drug degradation in a DP may be a function of water activity of the DP,[103] which can be kinetically controlled using packaging strategies such as the use of desiccants.

These studies highlight the need to quantitatively understand the variability in excipients and holistically address their impact on DP QAs taking into account the mechanism and kinetics of interactions between excipients, other formulation ingredients, packaging components, and DP manufacturing process. A robust control strategy can then be developed with the DP focus. Development of a control strategy for DPs and excipients to address potential effects of excipient variability needs to take into account the interaction of process with the excipient attributes.

Establishing tight specification controls on excipients is inherently challenging. A majority of pharmaceutical excipients are supplied by the food and chemical industry.[129] Excipients not only originate from diverse sources and constitute diverse material types (liquids, solids, semisolids, and gases), they are also used in diverse dosage forms and routes of administration. The varied uses create different performance expectations for each excipient. Compendial specifications for excipients are focused on safety and purity. With the increasing recognition of the functionality-related tests and the variability in excipients, the pharmacopeias have started to add elements of functionality testing to the excipients. As discussed earlier, the Ph.Eur. has included an FRC section to certain monographs, although compliance with these FRCs is not mandatory. The USP has included an Excipient Performance Chapter (<1059>) as an informational monograph that provides an overview of typical MAs associated with certain functional categories of excipients.[130] Moreton has published a series of articles discussing the functionality and performance of excipients in a QbD world[124,131,132] that address excipient attributes from the perspectives of excipient manufacturers. These articles highlight practical considerations that play a role in the design of excipient control strategies and the options available to the pharmaceutical scientists in understanding and incorporating excipient attribute variability within DP formulation design and development. Moreton and Carlin affirmed that tighter raw-material control is unlikely to be viable in practice and propose strategies such as robust formulation design and end-point adjustments within the manufacturing process that encompass the natural range of variability expected within excipients.[133]

6.5 RISK ASSESSMENT OF DRUG-EXCIPIENT INCOMPATIBILITIES AND MITIGATION STRATEGIES

Drug-excipient interactions can complicate and compromise a development program or the viability of a commercial product. Many of these interactions can take a long time to manifest and are not always predicted by stress and preformulation studies. However,

it is possible to reduce the probability of occurrence of such undesirable and costly scenarios by a careful risk assessment. This should be based on sound knowledge of the degradation pathways of the drug. A good awareness of excipient reactivity based on knowledge of reactive impurities that takes into account the nature, variability, and fate of these impurities is important. A well-designed and -executed drug-excipient compatibility study based on API properties such as reactivity and solid form, environmental factors of pH and moisture, formulation, and process design is essential. Early identification of drug-excipient incompatibilities followed by designing a DP that can tolerate excipient variability is the best approach to avoid undesirable surprises late in development and commercialization of the product.

There are several mitigation strategies one can adopt on a case-by-case basis to address the impact of drug-excipient incompatibilities. The reactive impurities in excipients may either be residues from excipient manufacturing process or degradation products of excipients and very often, the manufacturing process used by excipient vendors is a trade secret. Additionally, the pharmaceutical industry is relatively a smaller customer base for these excipients relative to the food and cosmetic industries, where demand is high and stringency requirements are relatively low. So the ability to influence the excipient vendor to modify the manufacturing process to suit the needs of pharmaceutical product manufacturers is limited requiring strong interaction. Changing the source of the excipient is another option, but it is often challenging. Where possible, avoiding the excipient is the best option. Many times that is not a viable option. In such situations, one can take other measures such as modifying the API form, formulation, and manufacturing process depending on where the issue resides. These can be through the use of stabilizers, pH modifiers that affect the microenvironmental pH, the solid form of the drug, or the formulation process, just to mention a few. Modifying packaging and storage conditions without modifying the formulation is another option. For example, desiccants such as silica gel can be copackaged with the dosage form. There are predictive models such as the sorption desorption moisture transfer model to simulate the effects of desiccants and determine the appropriate level of desiccant usage.

When changes to the formulation and packaging cannot be made, setting acceptance criteria for incoming material is another approach because it is not reasonable to expect compendial limits to address product-specific requirements. This can be done by an understanding of how changes in the reactive impurity level affects DP stability. There are challenges to this as well—for example, obtaining excipient samples with different levels of impurities from vendors is not easy. Analytical methods of these trace level impurities are another challenge. Finally, there are the cost considerations.

6.6 CONCLUSIONS

This chapter highlighted the cornerstones of excipient compatibility, functionality, and variability. Excipient compatibility studies are conducted with the primary goal of selecting dosage form components that are compatible with the drug. Methodically conducted experiments also provide additional information on stability profile of the drug, identify degradation products and mechanisms. Furthermore, if the stability of the drug is found to be lacking, strategies to mitigate the instability of the drug can be adopted. The guidelines and principles presented in this chapter would be useful in the appropriate design, conduct, and interpretation of compatibility studies to help accelerate formulation development activities and prevent or minimize surprises in drug development. Additionally, a clear understanding of excipient functionality in a DP, the underlying material properties that dictate the desired functionality, and the variability encountered in common excipients is expected to further enable the efficient development of DPs with sound control strategy and scientifically based specifications.

References

1. Mahato RI, Narang AS. *Pharmaceutical dosage forms and drug delivery*. 2nd ed. Boca Raton, FL: CRC Press; 2012. p. 472.
2. Carlin B, Moreton C. Flexible approaches to accommodate excipient variability using the principles of QbD. *Pharm Technol* 2010;**34**:S36–9.
3. Moreton RC. Functionality and performance of excipients in a quality-by-design world: excipient variability, QbD, and robust formulations. *Am Pharm Rev* 2010;**12**:24–7.
4. Hlinak AJ, Kuriyan K, Morris KR, Reklaitis GV, Basu PK. Understanding critical material properties for solid dosage form design. *J Pharm Innov* 2006;**1**:12–17.
5. Yu LX, Amidon G, Khan MA, Hoag SW, Polli J, Raju GK, et al. Understanding pharmaceutical quality by design. *AAPS J* 2014;**16**(4):771–83.
6. Badawy SI, Narang AS, LaMarche K, Subramanian G, Varia SA. Mechanistic basis for the effects of process parameters on quality attributes in high shear wet granulation. *Int J Pharm* 2012;**439**(1–2):324–33.
7. Badawy SIF, Narang AS, LaMarche KR, Subramanian GA, Varia SA, Lin J, et al. Integrated application of quality-by-design principles to drug product development: case study of brivanib alaninate film coated tablets. *J Pharm Sci* 2016;**105**(1):168–81.
8. Narang AS, Yamniuk A, Zhang L, Comezoglu SN, Bindra DS, Varia S, et al. Reversible and pH-dependent weak drug-excipient binding does not affect oral bioavailability of high dose drugs. *J Pharm Pharmacol* 2012;**64**(4):553–65.

9. Hickey MB, Peterson ML, Manas ES, Alvarez J, Haeffner F, Almarsson O. Hydrates and solid-state reactivity: a survey of b-lactam antibiotics. *J Pharm Sci* 2007;**96**(5):1090–9.
10. Rowe RC, Sheskey PJ, Weller PJ, Rowe R, Sheskey P, Weller P. *Handbook of pharmaceutical excipients*. 4th ed. Washington, DC: APhA Publications; 2003.
11. Kontny MJ. Distribution of water in solid pharmaceutical systems. *Drug Dev Ind Pharm* 1988;**14**(14):1991–2027.
12. Patel NK, Patel IJ, Cutie AJ, Wadke DA, Monkhouse DC, Reier GE. The effect of selected direct compression excipients on the stability of aspirin as a model hydrolyzable drug. *Drug Dev Ind Pharm* 1988;**14**(1):77–98.
13. Gore AY, Banker GS. Surface chemistry of colloidal silica and a possible application to stabilize aspirin in solid matrixes. *J Pharm Sci* 1979;**68**(2):197–202.
14. De Ritter E, Magid L, Osadca M, Rubin SH. Effect of silica gel on stability and biological availability of ascorbic acid. *J Pharm Sci* 1970;**59**(2):229–32.
15. Perrier PR, Kesselring UW. Quantitative assessment of the effect of some excipients on nitrazepam stability in binary powder mixtures. *J Pharm Sci* 1983;**72**(9):1072–4.
16. Fielden KE, Newton JM, O'Brien P, Rowe RC. Thermal studies on the interaction of water and microcrystalline cellulose. *J Pharm Pharmacol* 1988;**40**(10):674–8.
17. Ahlneck C, Alderborn G. Solid state stability of acetylsalicylic acid in binary mixtures with microcrystalline and microfine cellulose. *Acta Pharm Suec* 1988;**25**(1):41–52.
18. Yoshioka S, Aso Y, Terao T. Effect of water mobility on drug hydrolysis rates in gelatin gels. *Pharm Res* 1992;**9**(5):607–12.
19. Aso Y, Sufang T, Yoshioka S, Kojima S. Amount of mobile water estimated from 2H spin-lattice relaxation time, and its effects on the stability of cephalothin in mixtures with pharmaceutical excipients. *Drug Stab* 1997;**1**(4):237–42.
20. Snider B, Liang P, Pearson N. Implementation of water-activity testing to replace Karl Fischer water testing for solid oral-dosage forms. *Pharm Technol* 2007;**31**(2):56, 58, 60, 62, 64–66, 68, 70–71.
21. Heidemann DR, Jarosz PJ. Preformulation studies involving moisture uptake in solid dosage forms. *Pharm Res* 1991;**8**(3):292–7.
22. Burghart W, Burghart K, Raneburger J. *Solid formulation of levothyroxine and/or liothyronine salts containing controlled amount of water for stability*. Patent: (Globopharm Pharmazeutische Produktions- und Handelsgesellschaft m.b.H., Austria); 2004.
23. Ahlneck C, Zografi G. The molecular basis of moisture effects on the physical and chemical stability of drugs in the solid state. *Int J Pharm* 1990;**62**(2–3):87–95.
24. Shalaev EY, Zografi G. How does residual water affect the solid-state degradation of drugs in the amorphous state? *J Pharm Sci* 1996;**85**(11):1137–41.
25. Hancock BC, Zografi G. The relationship between the glass transition temperature and the water content of amorphous pharmaceutical solids. *Pharm Res* 1994;**11**(4):471–7.
26. Tong P, Zografi G. Effects of water vapor absorption on the physical and chemical stability of amorphous sodium indomethacin. *AAPS PharmSciTech* 2004;**5**(2):e26.
27. Chen JG, Markovitz DA, Yang AY, Rabel SR, Pang J, Dolinsky O, et al. Degradation of a fluoropyridinyl drug in capsule formulation: degradant identification, proposed degradation mechanism, and formulation optimization. *Pharm Dev Technol* 2000;**5**(4):561–70.
28. Glombitza BW, Oelkrug D, Schmidt PC. Surface acidity of solid pharmaceutical excipients. Part 1. Determination of the surface acidity. *Eur J Pharm Biopharm* 1994;**40**(5):289–93.
29. Stanisz B. The influence of pharmaceutical excipients on quinapril hydrochloride stability. *Acta Pol Pharm* 2005;**62**(3):189–93.
30. Crowley P, Martini L. Drug-excipient interactions. *Pharm Technol Eur* 2001;**13**(3):26–8 30–32, 34.
31. Janicki CA, Almond Jr HR. Reaction of haloperidol with 5-(hydroxymethyl)-2-furfuraldehyde, an impurity in anhydrous lactose. *J Pharm Sci* 1974;**63**(1):41–3.
32. Killion RB, Stella VJ. The nucleophilicity of dextrose, sucrose, sorbitol, and mannitol with p-nitrophenyl esters in aqueous solution. *Int J Pharm* 1990;**66**(1–3):149–55.
33. Badawy SI, Williams RC, Gilbert DL. Effect of different acids on solid-state stability of an ester prodrug of a IIb/IIIa glycoprotein receptor antagonist. *Pharm Dev Technol* 1999;**4**(3):325–31.
34. George RC. Investigation into the yellowing on aging of Sabril tablet cores. *Drug Dev Ind Pharm* 1994;**20**(19):3023–32.
35. Rivera SL, Ghodbane S. In vitro adsorption-desorption of famotidine on microcrystalline cellulose. *Int J Pharm* 1994;**108**(1):31–8.
36. Siepmann JI, Martin MG, Chen B, Stubbs JM, Potoff JJ. Monte Carlo simulations of supercritical fluid extraction systems. Abstracts of papers, 220th ACS national meeting, Washington, DC, August 20–24, 2000:COMP-143.
37. del Barrio MA, Hu J, Zhou P, Cauchon N. Simultaneous determination of formic acid and formaldehyde in pharmaceutical excipients using headspace GC/MS. *J Pharm Biomed Anal* 2006;**41**(3):738–43.
38. Hartauer KJ, Arbuthnot GN, Baertschi SW, Johnson RA, Luke WD, Pearson NG, et al. Influence of peroxide impurities in povidone and crospovidone on the stability of raloxifene hydrochloride in tablets: identification and control of an oxidative degradation product. *Pharm Dev Technol* 2000;**5**(3):303–10.
39. D'Souza AJ, Schowen RL, Borchardt RT, Salsbury JS, Munson EJ, Topp EM. Reaction of a peptide with polyvinylpyrrolidone in the solid state. *J Pharm Sci* 2003;**92**(3):585–93.
40. Chien YW, Van Nostrand P, Hurwitz AR, Shami EG. Drug--disintegrant interactions: binding of oxymorphone derivatives. *J Pharm Sci* 1981;**70**(6):709–10.
41. Cory W, Field K, Wu-Linhares D. Is it the method or the process-separating the causes of low recovery. *Drug Dev Ind Pharm* 2004;**30**(8):891–9.
42. Rohrs BR, Thamann TJ, Gao P, Stelzer DJ, Bergren MS, Chao RS. Tablet dissolution affected by a moisture mediated solid-state interaction between drug and disintegrant. *Pharm Res* 1999;**16**(12):1850–6.
43. Claudius JS, Neau SH. The solution stability of vancomycin in the presence and absence of sodium carboxymethyl starch. *Int J Pharm* 1998;**168**(1):41–8.
44. Senderoff RI, Mahjour M, Radebaugh GW. Characterization of adsorption behavior by solid dosage form excipients in formulation development. *Int J Pharm* 1982;**83**(1–3):65–72.
45. Lessen T, Zhao DC. Interactions between drug substances and excipients. 1. Fluorescence and HPLC studies of triazolophthalazine derivatives from hydralazine hydrochloride and starch. *J Pharm Sci* 1996;**85**(3):326–9.
46. Desai DS, Rubitski BA, Bergum JS, Varia SA. Effects of different types of lactose and disintegrant on dissolution stability of hydrochlorothiazide capsule formulations. *Int J Pharm* 1994;**110**(3):257–65.
47. Al-Nimry SS, Assaf SM, Jalal IM, Najib NM. Adsorption of ketotifen onto some pharmaceutical excipients. *Int J Pharm* 1997;**149**(1):115–21.
48. Johansen H, Moller N. Solvent deposition of drugs on excipients II: interpretation of dissolution, adsorption, and absorption characteristics of drugs. *Arch Pharm Chem (Sci)* 1977;**5**:33–42.

49. Tischinger-Wagner H, Endres W, Rupprecht H, Weingart A. [Oxidative degradation of linoleic acid methyl ester in suspensions of inorganic excipients. 1. Auto-oxidation in the presence of silicic acid products and aluminum oxide]. *Pharmazie* 1987;**42**(5):320–4.
50. Botha SA, Lotter AP. Compatibility study between naproxen and tablet excipients using differential scanning calorimetry. *Drug Dev Ind Pharm* 1990;**16**:673–83.
51. Rowe RC, Forse SF. Pitting: a defect on film coated tablets. *Int J Pharm* 1983;**17**(2–3):347–9.
52. Swaminathan V, Kildsig DO. An examination of the moisture sorption characteristics of commercial magnesium stearate. *AAPS PharmSciTech* 2001;**2**(4):28.
53. Kararli TT, Needham TE, Seul CJ, Finnegan PM. Solid-state interaction of magnesium oxide and ibuprofen to form a salt. *Pharm Res* 1989;**6**(9):804–8.
54. Ahlneck C, Waltersson J-O, Lundgren P. Difference in effect of powdered and granular magnesium stearate on the solid state stability of acetylsalicylic acid. *Acta Pharm Technol* 1987;**33**(1):21–6.
55. Thakur AB, Morris K, Grosso JA, Himes K, Thottathil JK, Jerzewski RL, et al. Mechanism and kinetics of metal ion-mediated degradation of fosinopril sodium. *Pharm Res* 1993;**10**(6):800–9.
56. Wirth DD, Baertschi SW, Johnson RA, Maple SR, Miller MS, Hallenbeck DK, et al. Maillard reaction of lactose and fluoxetine hydrochloride, a secondary amine. *J Pharm Sci* 1998;**87**(1):31–9.
57. Hemenway JN, Carvalho TC, Rao VM, Wu Y, Levons JK, Narang AS, et al. Formation of reactive impurities in aqueous and neat polyethylene glycol 400 and effects of antioxidants and oxidation inducers. *J Pharm Sci* 2012;**101**(9):3305–18.
58. Waterman KC, Arikpo WB, Fergione MB, Graul TW, Johnson BA, Macdonald BC, et al. N-methylation and N-formylation of a secondary amine drug (varenicline) in an osmotic tablet. *J Pharm Sci* 2008;**97**(4):1499–507.
59. Schildcrout SA, Risley DS, Kleemann RL. Drug-excipient interactions of seproxetine maleate hemi-hydrate: isothermal stress methods. *Drug Dev Ind Pharm* 1993;**19**(10):1113–30.
60. Huang T, Garceau ME, Gao P. Liquid chromatographic determination of residual hydrogen peroxide in pharmaceutical excipients using platinum and wired enzyme electrodes. *J Pharm Biomed Anal* 2003;**31**(6):1203–10.
61. Narang AS, Rao VM, Desai DS. Effect of antioxidants and silicates on peroxides in povidone. *J Pharm Sci* 2012;**101**(1):127–39.
62. Jansen PJ, Oren PL, Kemp CA, Maple SR, Baertschi SW. Characterization of impurities formed by interaction of duloxetine HCl with enteric polymers hydroxypropyl methylcellulose acetate succinate and hydroxypropyl methylcellulose phthalate. *J Pharm Sci* 1998;**87**(1):81–5.
63. Brewster ME, Loftsson T. The use of chemically modified cyclodextrins in the development of formulations for chemical delivery systems. *Pharmazie* 2002;**57**(2):94–101.
64. Li J, Guo Y, Zografi G. The solid-state stability of amorphous quinapril in the presence of beta-cyclodextrins. *J Pharm Sci* 2002;**91**(1):229–43.
65. Matilainen L, Larsen KL, Wimmer R, Keski-Rahkonen P, Auriola S, Jarvinen T, et al. The effect of cyclodextrins on chemical and physical stability of glucagon and characterization of glucagon/gamma-CD inclusion complexes. *J Pharm Sci* 2007;**97**(7):2720–9.
66. Eichman ML. *Drug-resin complexes stabilized by chelating agents*. Patent: Medeva Pharmaceuticals Manufacturing; 1999.
67. Bax RP. *Use of a combination of amoxycillin and clavulanate in the manufacture of a medicament for the treatment drug-resistant Streptococcus pneumoniae*. Patent: GlaxoSmithKline; 1998.
68. Crowley PJ. Excipients as stabilizers. *Pharm Sci Technol Today* 1999;**2**(6):237–43.
69. Wyttenbach N, Birringer C, Alsenz J, Kuentz M. Drug-excipient compatibility testing using a high-throughput approach and statistical design. *Pharm Dev Technol* 2005;**10**(4):499–505.
70. Plackett RL, Burman JP. The design of optimum multifactorial experiments. *Biometrika* 1946;**33**:305–25.
71. Durig T, Fassihi AR. Identification of stabilizing and destabilizing effects of excipient-drug interactions in solid-dosage form design. *Int J Pharm* 1993;**97**:161–70.
72. Cotton ML, Lamarche P, Motola S, Vadas EB. L-649,923—the selection of an appropriate salt form and preparation of a stable oral formulation. *Int J Pharm* 1994;**109**(3):237–49.
73. Landin M, Casalderrey M, Martinez-Pacheco R, Gómez-Amoza JL, Souto C, Concheiro A, et al. Chemical stability of acetylsalicylic acid in tablets prepared with different particle size fractions of a commercial brand of dicalcium phosphate dihydrate. *Int J Pharm* 1995;**123**(1):143–4.
74. Hancock BC, Zografi G. Characteristics and significance of the amorphous state in pharmaceutical systems. *J Pharm Sci* 1997;**86**(1):1–12.
75. Puttipipatkhachorn S, Yonemochi E, Oguchi T, Yamamoto K, Nakai Y. Effect of grinding on dehydration of crystal water of theophylline. *Chem Pharm Bull (Tokyo)* 1990;**38**(8):2233–6.
76. Takahashi Y, Nakashima K, Nakagawa H, Sugimoto I. Effects of grinding and drying on the solid-state stability of ampicillin trihydrate. *Chem Pharm Bull (Tokyo)* 1984;**32**(12):4963–70.
77. Nakagawa H, Takahashi Y, Sugimoto I. The effects of grinding and drying on the solid state stability of sodium prasterone sulfate. *Chem Pharm Bull (Tokyo)* 1982;**30**(1):242–8.
78. Kitamura S, Miyamae A, Koda S, Morimoto Y. Effect of grinding on the solid-state stability of cefixime trihydrate. *Int J Pharm* 1989;**56**(2):125–34.
79. Carlson E, Chandler W, Galdo I, Kudla T, Ta C. Automated integrated forced degradation and drug-excipient compatibility studies. *JALA* 2005;**10**(6):374–80.
80. van Dooren AA. Design for drug-excipient interaction studies. *Drug Dev Ind Pharm* 1983;**9**(1&2):43–55.
81. Sims JL, Carreira JA, Carrier DJ, Crabtree SR, Easton L, Hancock SA, et al. A new approach to accelerated drug-excipient compatibility testing. *Pharm Dev Technol* 2003;**8**(2):119–26.
82. Waltersson JO, Lundgren P. The effect of mechanical comminution on drug stability. *Acta Pharm Suec* 1985;**22**(5):291–300.
83. Badawy SI, Williams RC, Gilbert DL. Chemical stability of an ester prodrug of a glycoprotein IIb/IIIa receptor antagonist in solid dosage forms. *J Pharm Sci* 1999;**88**(4):428–33.
84. Guo Y, Byrn SR, Zografi G. Physical characteristics and chemical degradation of amorphous quinapril hydrochloride. *J Pharm Sci* 2000;**89**(1):128–43.
85. Hong J, Lee E, Carter JC, Masse JA, Oksanen DA. Antioxidant-accelerated oxidative degradation: a case study of transition metal ion catalyzed oxidation in formulation. *Pharm Dev Technol* 2004;**9**(2):171–9.
86. Katsumoto K, Ching TY, Theard LP, Current SP. *Multi-component oxygen scavenger system useful in film packaging*. Patent: US 5, 660, 761; 1997.
87. Huang T, Gao P, Hageman MJ. Rapid screening of antioxidants in pharmaceutical formulation development using cyclic voltammetry--potential and limitations. *Curr Drug Discov Technol* 2004;**1**(2):173–9.

88. Castello RA, Mattocks AM. Discoloration of tablets containing amines and lactose. *J Pharm Sci* 1962;**51**:106–8.
89. Cavatur R, Vemuri NM, Chrzan Z. Use of isothermal microcalorimetry in pharmaceutical preformulation studies. Part III. Evaluation of excipient compatibility of a new chemical entity. *J Therm Anal Calorim* 2004;**78**:63–72.
90. Giron D. Applications of thermal analysis and coupled techniques in pharmaceutical industry. *J Therm Anal Calorim* 2002;**68**:335–57.
91. Verma RK, Garg S. Selection of excipients for extended release formulations of glipizide through drug-excipient compatibility testing. *J Pharm Biomed Anal* 2005;**38**(4):633–44.
92. Schmitt E, Peck K, Sun Y, Geoffroy J. Rapid, practical, and predictive excipient compatibility screening using isothermal microcalorimetry. *Thermochim Acta* 2001;**380**:175–83.
93. Selzer T, Radau M, Kreuter J. The use of isothermal heat conduction microcalorimetry to evaluate drug stability in tablets. *Int J Pharm* 1999;**184**(2):199–206.
94. Selzer T, Radau M, Kreuter J. Use of isothermal heat conduction microcalorimetry to evaluate stability and excipient compatibility of a solid drug. *Int J Pharm* 1998;**171**(2):227–41.
95. Zhang GG, Law D, Schmitt EA, Qiu Y. Phase transformation considerations during process development and manufacture of solid oral dosage forms. *Adv Drug Deliv Rev* 2004;**56**(3):371–90.
96. Morris KR, Griesser UJ, Eckhardt CJ, Stowell JG. Theoretical approaches to physical transformations of active pharmaceutical ingredients during manufacturing processes. *Adv Drug Deliv Rev* 2001;**48**(1):91–114.
97. Airaksinen S, Karjalainen M, Kivikero N, Westermarck S, Shevchenko A, Rantanen J, et al. Excipient selection can significantly affect solid-state phase transformation in formulation during wet granulation. *AAPS PharmSciTech* 2005;**6**(2):E311–22.
98. Narang AS, Desai D, Badawy S. Impact of excipient interactions on solid dosage form stability. *Pharm Res* 2012;**29**:2660–83.
99. Panakanti R, Narang AS. Impact of excipient interactions on drug bioavailability from solid dosage forms. *Pharma Res* 2012;**29**(10):2639–59.
100. Narang AS, Desai D, Badawy S. Physicochemical interactions in solid dosage forms. *Pharm Res* 2012;**29**(10):2635–8.
101. Wu Y, Levons J, Narang AS, Raghavan K, Rao VM. Reactive impurities in excipients: profiling, identification and mitigation of drug-excipient incompatibility. *AAPS PharmSciTech* 2011;**12**(4):1248–63.
102. Wasylaschuk WR, Harmon PA, Wagner G, Harman AB, Templeton AC, Xu H, et al. Evaluation of hydroperoxides in common pharmaceutical excipients. *J Pharm Sci* 2007;**96**(1):106–16.
103. Hemenway J, Kirby S, Narang A, Rao V, Paruchuri S, Derbin G, et al. Effect of water activity and water absorption properties on the stability of film-coated tablet formulations of a moisture sensitive active pharmaceutical ingredient. In: *American Association of Pharmaceutical Scientists annual meeting*, New Orleans, LA; 2010.
104. Narang AS, LaMarche K, Subramanian G, Lin J, Varia S, Badawy S. Granule porosity is the overriding mechanism controlling dissolution rate of a model insoluble drug from an immediate release tablet. In: *American Association of Pharmaceutical Sciences*, Washington, DC; 2011.
105. Wolfe S, Pafiakis S, Tang D, Jennings S, Abebe A, Gao Z, et al. Effect of tablet pore size distribution on dissolution upon accelerated storage conditions. In: *AAPS annual meeting*, San Antonio, TX; 2013.
106. Narang AS. Excipients impacting bioavailability: a case for simplicity in formulation. Website: <http://aapsblog.aaps.org/>; 2013.
107. Rees JE, Rue PJ. Interbatch variation in compression properties of Elcema G250 [proceedings]. *J Pharm Pharmacol* 1978;**30**(Suppl):25P.
108. Butcher AE, Jones TM. Some physical characteristics of magnesium stearate. *J Pharma Pharmacol* 1972;**24**(Suppl):1P–9P.
109. Holzer AW, Sjogren J. Evaluation of some lubricants by the comparison of friction coefficients and tablet properties. *Acta Pharm Suec* 1981;**18**(3):139–48.
110. Barra J, Somma R. Influence of the physicochemical variability of magnesium stearate on its lubricant properties: possible solutions. *Drug Dev Ind Pharm* 1996;**22**(11):1105–20.
111. Chamarthy SP, Pinal R, Carvajal MT. Elucidating raw material variability--importance of surface properties and functionality in pharmaceutical powders. *AAPS PharmSciTech* 2009;**10**(3):780–8.
112. Williams III RO, Sriwongjanya M, Barron MK. Compaction properties of microcrystalline cellulose using tableting indexes. *Drug Dev Ind Pharm* 1997;**23**(7):695–704.
113. Narang AS. Addressing excipient variability in formulation design and drug development. In: Narang AS, Boddu SH, editors. *Excipient applications in formulation design and drug delivery*. Springer International Publishing; 2015. p. 541–68.
114. Kushner IV J, Langdon BA, Hiller JI, Carlson GT. Examining the impact of excipient material property variation on drug product quality attributes: a quality-by-design study for a roller compacted, immediate release tablet. *J Pharm Sci* 2011;**100**(6):2222–39.
115. Kushner JT. Utilizing quantitative certificate of analysis data to assess the amount of excipient lot-to-lot variability sampled during drug product development. *Pharma Dev Technol* 2013;**18**(2):333–42.
116. Kushner J. Utilizing quantitative certificate of analysis data to assess the amount of excipient lot-to-lot variability sampled during the development of new drug products. In: *AAPS annual meeting and exposition*, New Orleans, LA; 2010.
117. Wu H, Tawakkul M, White M, Khan MA. Quality-by-design (QbD): an integrated multivariate approach for the component quantification in powder blends. *Int J Pharm* 2009;**372**(1–2):39–48.
118. Alvarez-Lorenzo C, Castro E, Gomez-Amoza JL, Martinez-Pacheco R, Souto C, Concheiro A. Intersupplier and Inter-lot variability in hydroxypropyl celluloses: implications for theophylline release from matrix tablets. *Pharm Acta Helv* 1998;**73**(2):113–20.
119. Doelker E, Mordier D, Iten H, Humbert-Droz P. Comparative tableting properties of sixteen microcrystalline celluloses. *Drug Dev Ind Pharm* 1987;**13**(9–11):1847–75.
120. L'Hore-Gaston J, Schmitt R, Li Y, Karas C. Using polyethylene oxide blends to assess the effect of excipient variability: in formulation development of hydrophilic extended-release matrix tablets. *Pharm Tech* 2009;**33**:36–41.
121. Carvalho T, Rao V, Levons J, Wu J, Narang A, Paruchuri S, et al. Formation of reactive impurities in PEG 400 and effects of antioxidants and oxidation inducers on the reaction rates. In: *American Association of Pharmaceutical Sciences annual meeting*, Washington, DC; 2011.
122. Gabrielsson J, Sjoestroem M, Lindberg N-O, Pihl A-C, Lundstedt T. Robustness testing of a tablet formulation using multivariate design. *Drug Dev Ind Pharm* 2006;**32**(3):297–307.
123. Kushner J, Langdon BA, Hicks I, Song D, Li F, Kathiria L, et al. A quality-by-design study for an immediate-release tablet platform: examining the relative impact of active pharmaceutical ingredient properties, processing methods, and excipient variability on drug product quality attributes. *J Pharm Sci* 2014;**103**(2):527–38.

124. Moreton C. Functionality and performance of excipients in quality-by-design world Part 4: obtaining information on excipient variability for formulation design space. *Am Pharm Rev* 2009;**12**(5):28, 31–33
125. Narang A, Kanthasamy M, Castoro J, Varia S, Desai D. Effect of pro- and anti-oxidants on the formation of formyl species in PVA- and PEG-based tablet coating material. In: *American Association of Pharmaceutical Sciences annual meeting*, Washington, DC; 2011.
126. Narang AS, Rao V, Farrell T, Ferrizzi D, Castoro J, Corredor C, et al. Stability implications of prolonged storage of PVA and PEG-based coating suspension. In: *American Association of Pharmaceutical Sciences annual meeting*, New Orleans, LA; 2010.
127. Reed RA. Excipient mediated oxidation by organo-peroxides. *Presentation at the American Association of Pharmaceutical Scientists (AAPS)*, Baltimore, MD; 2004.
128. Reed RA, Harmon P, Manas D, Wasylaschuk W, Galli C, Biddell R, et al. The role of excipients and package components in the photostability of liquid formulations. *PDA J Pharm Sci Technol/PDA* 2003;**57**(5):351–68.
129. Dave VS, Saoji SD, Raut NA, Haware RV. Excipient variability and its impact on dosage form functionality. *J Pharm Sci* 2015;**104**(3):906–15.
130. Sheehan C, Amidon GE. Compendial standards and excipient performance in the QbD era: USP excipient performance chapter <1059>. *Am Pharm Rev* 2011;**14**(6):10 12, 14–16, 18.
131. Moreton C. Functionality and performance of excipients in quality-by-design world part VI: excipient composition. *Am Pharm Rev* 2009;**12**(7):48–51.
132. Moreton C. Functionality and performance of excipients in a quality-by-design world part 2: excipient variability, QbD and robust formulations. *Am Pharm Rev* 2009;**12**(2):40 42–43, 45.
133. Moreton RC, Carlin B. Flexible approaches to accommodate excipient variability using the principles of QbD. *PharmTech* 2010;**34**:1–4.

This page intentionally left blank

CHAPTER

7

Polymer Properties and Characterization

J. Brady[1], *T. Dürig*[1], *P.I. Lee*[2] *and J.-X. Li*[3]

[1]Ashland Inc., Wilmington, DE, United States [2]University of Toronto, Toronto, ON, Canada
[3]Ferring Pharmaceuticals, Parsippany, NJ, United States

7.1 INTRODUCTION

The introduction of polymer excipients has been crucial in advancing modern formulation design and solid oral drug delivery technologies over the past 50 years. Polymeric materials are widely used in a broad range of pharmaceutical products and constitute essential components of modern solid oral dosage forms. Therefore, it is important to understand polymer properties and characterization methods in order to enable rational design and development of oral solid drug delivery systems and manufacturing processes.

The word "polymer" is derived from the Greek πολυ *(polu)* for "many" and μέρος *(meros)* for "part," meaning a long-chain consisting of many parts. The combination of repeating units to form long chains with widely variable length and structure is essential in determining the properties and applications of the resulting polymers, as the diversity of polymer properties is attributed to the simple fact that each polymer molecule consists of many parts with individual variations in sequence of repeating units and chain length.

In contrast to substances comprised of small molecules that are well defined and identified by a set of system variables, including basic physical properties, states, or characteristics with unique and discrete values, polymers are less well defined or characterized in the conventional sense. This applies to even the simplest of polymeric systems. For example, substances comprised of small molecules have a distinct molecular weight, with each molecule having exactly the same nominal mass, neglecting, for the moment, the multiple natural isotopes possible for the atomic constituents. Polymers, on the other hand, are characterized by a diverse population in which there is a broad distribution of molar masses across the entire chain population. This is well known and clearly understood, since polymers are simply chains of many monomeric residues covalently linked together. For small molecules, it can be clearly defined if a molecule is in solution or not. Such a clear distinction may not be applied to polymers. It is quite possible for extended segments of a long polymer chain to effectively reside in bulk solution, while other parts still reside in a bulk polymer solid phase. Conceptually, the polymer chain is not completely in solution, yet the behavior of this system is clearly very different from the one where an unswollen polymer sample is sitting unsolvated in a sea of solvent. The state of an unswollen polymer is not a fluid or solution by traditional definition. Distributions in properties, states, and characteristics are what set the behavior of many polymers apart from small molecules in general. Understanding the operational impact of this difference is the first step in rational characterization and applications of polymeric materials.

This chapter will emphasize some of the pragmatic consequences of continuous distributions in composition, molecular weight, and molecule state in the use of polymeric materials and the impact on their functionality. Since the basic literature of polymer chemistry is readily available,[1–6] we will strive not to duplicate that information here. Rather, this chapter should be viewed as an operational entry point into the practical utilization of polymeric materials. While we will attempt to maintain a firm foot in fundamental science, each of the major subsections of this chapter could constitute a monograph in its own right. As such, the approach taken is to cover the information that will be of immediate value to a user of polymeric excipients. This necessitates a somewhat limited but focused scope of coverage. In addition, we do not wish

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00007-8

to be distracted by equations that only approximately describe the major features of polymeric systems and often obscure important underlying operational principles. Where possible, we will resort to experimental rather than theoretical examples to illustrate specific points. Our objective is to provide the user and formulator of polymeric materials an operational and pragmatic guide upon which to base formulation decisions and understand the molecular origins of these performance metrics.

In this chapter, we will consider basic structural issues surrounding polymeric materials. One general aspect to appreciate is that two samples with identical average molecular weight and composition may display very different behavior in specific applications. Since a polymer is comprised of chains of monomer units, the total number of monomers per chain (ie, the molecular weight), the gross topology of the chain (linear vs branched), and the microstructural monomer sequence within and across the chains (ie, block A−B copolymer vs random A−B copolymer vs a physical mixture of polymers A and B, see Fig. 7.1) all impact the final properties of the material.

Determination of the average monomer composition and molecular weight of a polymeric sample are experiments routinely carried out using a variety of well-known techniques. Other attributes of polymeric materials, for example, branching or large-scale composition heterogeneity across chains, may require some level of experimental finesse but can be addressed in broad terms. However, detailed questions regarding monomer sequencing within chains can be all but impenetrable in most instances and often require appeal to indirect challenge/response using experimental probes to even qualitatively address them. This fine detail is of significant importance in current research, since key performance properties can often be directly related to this elusive structural parameter.

7.1.1 Definition, structure, and nomenclature

The gross architecture of a polymer chain can be readily assessed by asking whether the chain or chain assembly can be mapped onto a one- (linear), two- (branched), or three- (cross-linked network) dimensional object (Fig. 7.2). In general terms, these distinctions reflect the natural spatial connectivity imparted by these chains when placed into a volume element, either in formal solution, a melt, or the solid state.

How a polymeric chain mechanically couples to and transports within a medium comprised of either other polymeric materials or low molecular weight solvent is determined in part by the gross topology of the chain system. Ultimately, the key physical attributes are the length, scale, and nature of connectivity between volume elements in solution, gel, melt, or solid states and how that connectivity propagates through space.

If one takes a linear chain and stretches it, a simple line is obtained. By the same token, the lowest

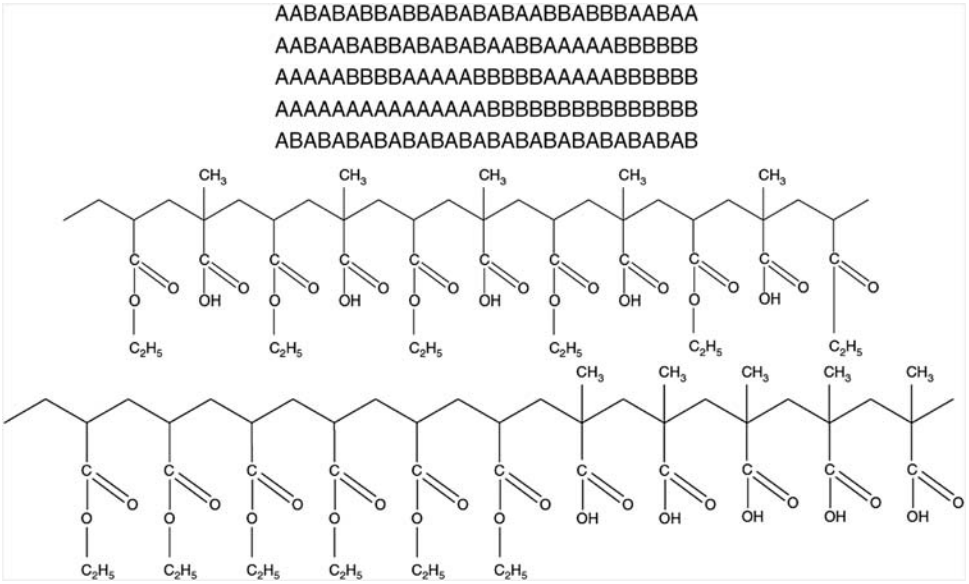


FIGURE 7.1 Examples of AB type copolymers in which the composition is the same (50:50 A:B), but the sequence of A and B monomers differ. Prototype copolymer structures shown top to bottom represent random, block graft, multiblock, diblock, and alternating AB copolymer systems. As a specific chemical example, consider the case of a 50:50 copolymer comprised of ethyl acrylate and acrylic acid as either an alternating, or diblock, architecture. The solution properties of these two materials, which have equal composition and molar mass, would be expected to be vastly different, as would solid state and melt properties.

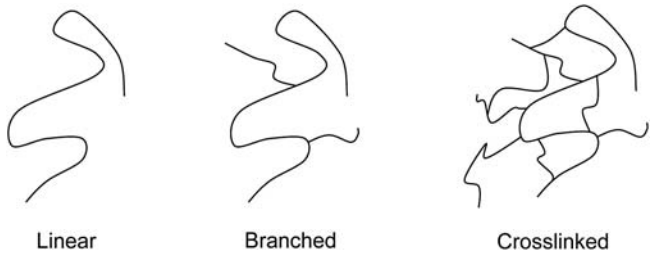


FIGURE 7.2 Diagrammatic examples of prototype chain architectures. A linear chain has been depicted as a random coil, although rigid or semirigid rods are also possible. The branched example is comprised of a main backbone with two extended long-chain branches, while more extensive branching and linking to other main backbones have been used to depict a typical cross-linked polymeric system.

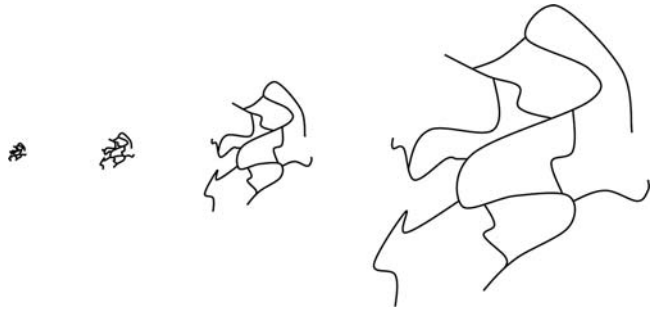


FIGURE 7.3 Depending upon the length and the flexibility of chain segments between covalent cross-links, a cross-linked polymer may be a very rigid and nonswelling solid (short, rigid segments between cross-links) or very highly swelling (low cross-link density between very flexible and solvent compatible segments) under the appropriate solvent conditions. In some cases, for example with a change of pH from highly acid to neutral or basic for a cross-linked polyacrylic acid system, conversion between collapsed and highly swollen states is readily and reversibly accomplished.

dimensional object that a simple branched structure will yield under deformation, in the simplest case, is a two-dimensional planar object. A cross-linked network will yield a three-dimensional volume element under any applied deformation. In this sense, these three classes of chains are topologically and dimensionally distinct. As in any discussion involving polymeric species, the evolution between types is continuous rather than discrete even when involving dimensional extent; simple characteristics, lengths, scales along principle axes are critical.

For example, as a simple linear chain is grafted with an increasingly long side chain in the midsection of the molecule, the evolution from a purely linear to a branched planar structure occurs in a continuous fashion. Likewise, as multiple chains having a two-dimensional architecture receive additional grafts on side chains, a similar transition from two- to three-dimensional character occurs. It is critical to appreciate that even aspects that appear firmly rooted in discrete representations, dimensionality in this example, may take on some attributes of a continuous variable if viewed from specific contexts.

The chain types shown in Fig. 7.2 are assumed to represent covalently linked permanent structures. Polymeric materials of a given intrinsic topology may also assume a higher apparent dimensionality via dynamic self-association. This association can involve either similar or compositionally distinct partners. Both types of examples exist in use, with self-association of a hydrophobically modified polymer[7] reflecting the former case, while the interaction of xanthan and locust bean gum or sodium carboxymethyl cellulose and hydroxyethyl cellulose are examples of the latter.[8] Finally, simple physical entanglements at high polymer concentrations can be viewed as a type of dynamic cross-link, since many of the essential properties of cross-linked systems (structural robustness, high elasticity, slow transport of polymer, and/or constituent polymer) will be achieved, particularly at relatively short timescales (minutes to hours).

Gross chain topology has an influence on the flow and mechanical properties of polymer melts and solutions, as well as over the timescales of structural relaxation. In general, as a system transitions from pure linear to highly cross-linked, the balance between the viscous and elastic nature of a solution or melt tends towards more elastic. This transition is similar in kind to that observed as the molecular weight of a polymer is increased. In terms of gross phenomenology, the basic causes are tied to similar origins, with the understanding that in a cross-linked or branched polymer sample permanent covalent linkages dominate the situations, while with high-molecular-weight linear polymers dynamic interchain entanglements, which can be akin to a transient cross-link, control system response.

In addition to extensively cross-linked pure polymer systems, an intermediate type of system is exemplified by cross-linked swellable microparticulates—the most common of these being provided by cross-linked acrylic acid-based systems prepared in a microparticulate form. These are often used to control the flow properties of systems by simply encapsulating substantial fractions of the total system volume but keeping it segregated into rather small length scale packets (see Figs. 7.3 and 7.4). With large swelling ratios, relatively small amounts of particulate mass can basically encapsulate large volumes of liquid. This is often useful to build in a yield stress while maintaining a relatively low-shear viscosity, since the propagation of shear forces is basically constrained to simple soft sphere hydrodynamic effects (vs chain entanglement for a molecular dissolved system), with a net length

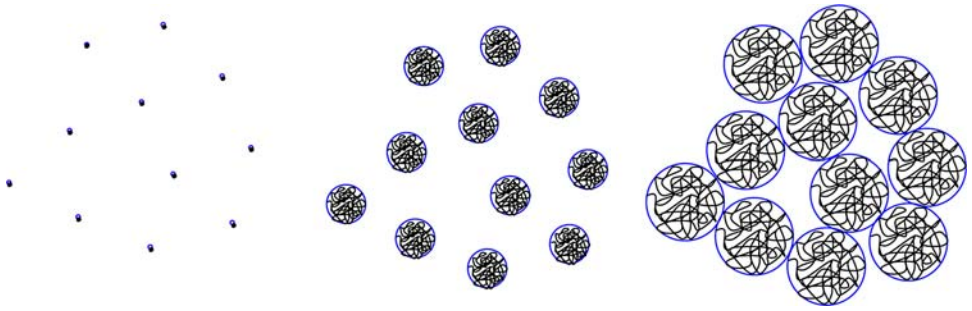


FIGURE 7.4 In addition to bulk cross-linked systems, the swelling of cross-linked microparticulate polymeric systems provides interesting opportunities for rheology control. As the particles swell, they first occupy a significant volume fraction of the dispersed phase and then reach a point at which all discrete swollen particles start to overlap and touch. Diameter ratios of 1:10:20, yielding a net 8000-fold volumetric swelling, is depicted.

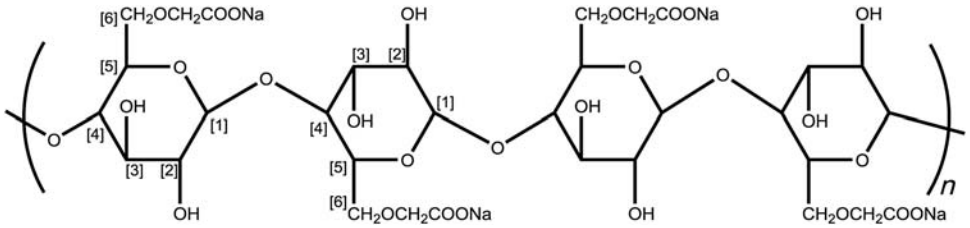


FIGURE 7.5 Representative structure of sodium carboxymethyl cellulose. For graphic convenience, only substitution at the six-position is explicitly shown. However, all three hydroxyl positions are available for substitution, and a real polymer will be comprised of a mixture of primarily unsubstituted, monosubstituted, or disubstituted monomer residues. This example is an idealized DS = 1 sodium carboxymethyl cellulose.

scale established by the dimensions of the swollen particle. In general, our attention will be restricted to linear polymers, since they comprise the majority of systems of direct interest.

7.1.2 Types of homopolymers and copolymers

In dealing with polymeric materials, cautious delineation of whether a material is a homopolymer or a more complex copolymer is an extremely critical initial step in understanding the range of properties that can emerge in everyday use. In general, most of the polymeric materials that we deal with are much more profitably handled and understood if it is explicitly recognized that they are copolymeric in reality, and the detailed sequence of the various residues is an important variable that can influence application properties in use. In some cases the distinctions are clear, while in others they are not.

As a specific example of the potentially significant impact that a sequence can have on a material, consider a simple AB copolymer that possesses equal molar amounts of A and B (Fig. 7.1). For the same average chemical composition, molecular weight, and gross architecture, a significant range of structural archetypes can be prepared by manipulating only the sequence of the monomer units. That sequence can run from purely random to somewhat blocky, from multiblock to diblock, and to purely alternating structures. If the monomer residues possess differing solution properties, for example consider a case in which one residue is relatively hydrophobic while the other is very hydrophilic, the sequence can have an overwhelming influence on properties such as solubility, phase behavior, and packing in the solid state. As a specific case, consider poly(*co*-ethyl acrylate-methacrylic acid) (Fig. 7.1). The alternating and diblock architectures will yield vastly different physical properties (solubility, mechanical properties) and performance traits despite the equivalence of molecular weight and average composition. Naturally, the argument can be extended to include a physical mixture of the two homopolymers, which provide the same average molecular weight and composition values. Again, one has a system of near-identical molecular weight and composition, but the performance is anticipated to be distinctly different. In fact, the component polymers would likely be immiscible in the melt state and prone to segregate if deposited from the solution state in, for example, a coating application.

As a second example, consider the simple case of sodium carboxymethylcellulose with an average degree of substitution (DS) of one. In total, carboxymethyl substitution can occur at up to three hydroxyl positions on any anhydroglucose residue (Fig. 7.5).

In principle, a sodium carboxymethylcellulose chain can be comprised of unsubstituted anhydroglucose, have monosubstitution at either the two, three, or six hydroxyl positions, disubstitution at the 2/3, 2/6, or 3/6 positions, or be a fully trisubstituted monomer. Even if the added detail of positional substitution is ignored and we consider only unsubstituted, monosubstituted, disubstituted, and trisubstituted monomers as distinct entities, the system has to be considered as an ABCD copolymer. As noted, the detailed sequence distribution of the A, B, C, and D monomer residues can have an impact on the final properties of the material if the intrinsic properties of those constituent monomers (polarity, hydrophobicity/hydrophilicity, charge state, propensity to participate in interchain hydrogen bond association) differ significantly. Obviously, the solution and solid-state behavior of a neutral sugar (anhydroglucose) will be different from a trianion salt (sodium tricarboxymethyl anhydroglucose) due to the ionic charge and steric packing requirements of the latter and the facile possibility for hydrogen bond formation for the former.

As illustrated by the examples, one expectation is that microstructural sequence variations will impact the local chemical nature of subsections within a polymer chain. In the case of poly(*co*-ethyl acrylate-methacrylic acid), we see a transition from a nominally hydrophilic alternating copolymer to a system with distinct hydrophobic and hydrophilic blocks. In terms of expected solubility in water, the hydrophobic block will not find an aqueous environment amenable. Depending on the net balance of segment lengths and position along the backbone, this may manifest itself as anything from partial insolubility with a high propensity to swell, complete solubility with a tendency to adsorb onto surfaces, a tendency to self-associate in solution, or as gross insolubility in pure solvents with a decided need to employ binary solvent mixtures to achieve complete dissolution. A reality with polymeric materials is that all of these possibilities can, in principle, be expressed by materials of the same average composition via adjustment of molecular weight and chain microstructure sequence.

The case with synthetic copolymers is fairly obvious and readily understood on the basis of the respective behavior of the constituent monomers. It is a less obvious situation with derivatized cellulosics, even though the underlying issues are identical and similar trait differences for nominally similar compositions are possible, although the underlying physical phenomena are somewhat distinct. With cellulosic derivatives, aside from the cases of cellulose esters and ethyl cellulose, the possible constituent monomers are all fairly hydrophilic. However, similar issues remain and tend to be tied to the potential formation of hydrogen bonding junction zones in structurally regular domains of the material. These structurally regular domains can be due to short runs of contiguous unsubstituted anhydroglucose or uniformly substituted contiguous run lengths. In both cases, the underlying feature is a structurally regular domain that can be annealed into a highly ordered domain for which the primary events of dissolution (solvent penetration and separation of the packed chains) are kinetically unfavorable due to cooperative multipoint hydrogen bond association.

This added level of complexity, which is intrinsic with derivatized cellulosic excipients but an important operational detail for many polymeric samples, needs to be appreciated in order to understand why materials that possess very similar average compositions can display rather diverse end use performance attributes in different applications.

7.2 BASIC CONCEPTS AND CHARACTERIZATION OF POLYMERIC MATERIALS

Some specific example polymers were described in Section 7.1, as were some of the key traits used to characterize these polymers. In this section we'll delve somewhat deeper into the measurement of these traits and the interrelationships between them.

At the most basic level, a polymeric material is defined by the average molecular weight distribution and average composition. As has already been seen with molecular weight, the net range in molecular weight sampled is enormous. This can have important consequences. Traits that are dependent on properties showing variation, take molecular weight as a specific example, may display particular sensitivity to specific portions of the molecular weight distribution. What this means is that those traits that are molecular-weight dependent may differ between materials with nominally the same average molecular weights. A specific example would be the complete shear rate dependent flow curve of a blend of high and low molecular weight materials with that which is a pure intermediate molecular weight.

7.2.1 Polymer composition

At the basic level, polymers are characterized by average composition variables that reflect either the total level of secondary substitution for derivatized systems (eg, cellulose ethers) or the level of constituent monomers for a synthetic or natural copolymer system.

For example, cellulosic derivatives are compositionally characterized by the percent weight of the functional group attached to the backbone, the degree of substitution (DS) per anhydroglucose, or the total molar substitution (MS) per anhydroglucose residue. These three modes of characterization are largely interchangeable with the preferred quantity often dependent on historical legacy and whether the substituent can form an oligomeric pendant group. In the latter case, the DS would not fully characterize the net amount of substituent on a chain.

It is important to recognize that the definition of an operational functional group may include contributions from both the anhydroglucose residue and the added functional reagent, leading to apparent percentage of weight function groups that may appear inordinately high. A good example of this would be ethyl cellulose, in which the average composition is typically expressed as the percentage of ethoxyl content. In this case, the ethyl portion of the ethoxyl group is derived from the derivatizing reagent (ethyl chloride), while the oxygen atom is provided by the anhydroglucose residue. For a typical N or standard type ethylcellulose, a DS of 2.5 corresponds to a weight percentage of ethoxyl of 48.5. However, the actual percentage of mass imparted by the added ethyl group is 31.3%, with the remainder contributed by the oxygen atom that is already a part of the main backbone polymer.

The second manner to characterize the composition of cellulosics is to speak of the DS. Each anhydroglucose residue along a cellulosic backbone possesses three reactive hydroxyl groups, located specifically at the C2, C3, and C6 positions. The average DS of a material quantifies the average number of hydroxyl groups that are derivatized per anhydroglucose. The maximum DS of any cellulosic material is 3, at which point all reactive hydroxyl functionality will have been consumed. The C2 and C6 positions are the most reactive sites and will generally carry the bulk of the substituents.

When the derivatization reagent effectively caps the reactive hydroxyl functionality (eg, methyl, ethyl, or carboxymethyl functionality), the amount of bound functionality quantifies both the degree of hydroxyl substitution and total MS. However, as already noted, if the substituent is able to oligomerize, as is the case with hydroxyethyl or hydroxypropyl substitution, the connection between DS and MS starts to diverge. Furthermore, in principle, the net MS provided by an oligomerizing functional group is not bound to an upper limit determined by the number of backbone reactive sites, although steric constraints and reaction efficiency do provide a pragmatic limit. Of all available cellulosics, hydroxypropyl cellulose (HPC) is the most highly substituted system, with MS levels routinely in the vicinity of four.

In the case of synthetic copolymers, composition is typically determined by the average mole percent or weight percent of the various copolymer constituents, and/or the charge state for a titratable functionality. In general, this is controlled by the charge composition of the reaction mixture used to prepare the product.

In either case, there are a variety of experimental techniques that can be used to quantify the mean composition of the material. These techniques include direct functional group analysis, in the case of the derivatized cellulosics, while the average monomer composition is generally the target for synthetic copolymer systems. In part, this is related to the potential complexity with cellulosics that possess chain-extended pendant groups.

One significant compositional difference between polymers and low-molecular-weight materials is that the detailed sequence distribution of the constituents at equivalent average compositions can have a significant impact on the final solution and solid-state properties of the polymer.

In contrast to the situation that exists with the determination of the average material composition, there is a dearth of direct experimental techniques that provide information on the sequence of monomers in a copolymer system. In general, one is left with fairly indirect experimental probes to assess monomer sequence.

In broad strokes, these indirect probes fall into two general classifications. First, there are the techniques that degrade the polymer to an oligomeric species and then analyze the compositional heterogeneity of these oligomeric materials. While this approach provides insight into the compositional heterogeneity of the system, ascribing the difference observed between interchain and intrachain sequence variations is sometimes quite ambiguous, since one is unable to trace the fragments back to the originating polymers.

The second strategy is to examine the response of the system to the applied challenges. Specific examples would include examining the solubility behavior in a variety of pure and mixed solvent systems, characterizing solubility as a function of temperature (cloud point) or using site-directed cleavage as an approach to probe specific structural motifs. An example of the latter would be enzymatic degradation, which is selective to the presence of short runs of contiguous unsubstituted anhydroglucose. Both approaches have appeared in the literature.[9–13]

In each case, an indirect response is used to provide qualitative insight into the presence or absence of specific run sequences. The utility of this approach is that it provides insight into the origins of observable performance differences between materials that have the same average composition and molecular weight.

At this stage, the information derived from these types of measurements remains semiquantitative.

7.2.2 Molecular weight

Owing to the dispersity in chain lengths and potentially number of substituents per residue of an assembly of polymeric molecules, molecular weights must be discussed in terms of distribution averages. Three distinct types of averages are generally considered: the number, the weight, and z-average molecular weights.

Defining relationships for these three quantities are as follows:

$$M_n = \frac{\sum_i n_i M_i}{\sum_i n_i} = \frac{\sum_i w_i}{\sum_i \frac{w_i}{M_i}}$$

$$M_w = \frac{\sum_i n_i M_i^2}{\sum_i n_i M_i} = \frac{\sum_i w_i M_i}{\sum_i w_i}$$

$$M_z = \frac{\sum_i n_i M_i^3}{\sum_i n_i M_i^2} = \frac{\sum_i w_i M_i^2}{\sum_i w_i M_i}$$

with

$$w_i = n_i M_i$$

where:

w_i is the mass of the fraction of polymer chains with molecular weight M_i;
n_i is the number of these chains.

For a perfectly monodisperse polymer sample, $M_w = M_n$. One more frequently speaks of the population polydispersity or M_w/M_n, which is equal to 1.0 for a perfectly monodisperse sample, since all the chains in that sample have precisely the same molecular weight. For the materials being discussed here, sample polydispersities will generally fall in the range of 10–20, dropping down to values of 3 or so for rather low-molecular-weight materials, and occasionally rising above 20 if the systems of differing viscosity grade have been blended. In understanding the performance of polymeric materials, high-molecular-weight polydispersity can often be a decided benefit. The most extreme example of this would be a synthetic polymer in which a residual monomer serves as an integrated plasticizer of the polymeric solid to help control mechanical properties of the system. To varying extents, most low-molecular-weight constituents (including moisture and low-molecular-weight oligomers) can serve this type of role.

While the three molecular weight averages may appear to be a simple mathematical construct to characterize a polydisperse population, the origin of these relations resides in the different classical methods that one can employ to experimentally determine the absolute molecular weight of a polymer and the distinct averaging schemes operative in these methods. The various techniques used to measure polymer molecular weight are well described in the literature.[14–16]

Colligative property measurements of absolute molecular weight (vapor pressure lowering, boiling point elevation, osmotic pressure) are sensitive to the number average molecular weight. Techniques such as light scattering can, under certain circumstances, yield the weight-average molecular weight. Finally, ultracentrifugation-based methods yield z-average molecular weight. In other words, the differing physical response for each of the measurement approaches provides intrinsically different forms of averaging within the polydisperse population. In addition to these absolute methods for determining molecular weight, any physical observable that possesses a molecular weight dependence, for example, solution viscosity at a fixed concentration can be used as an indirect proxy for a molecular weight determination, with operational correlation methods providing the needed link to the actual molecular weight.

In recent years, size exclusion chromatography (SEC) has largely replaced these classical approaches for routine absolute molecular weight determination. SEC has a distinct advantage over the classical measurements in that it is able to provide direct information on the complete molecular weight distribution via a direct separation based on the hydrodynamic size of the molecules in solution. There are some additional pragmatic issues that potentially improve the robustness of the SEC-based measurement including:

- Since a separation is performed, contamination of the sample with low-molecular-weight impurities has an inconsequential impact on the final result, since their contribution can be ignored at the data processing step.
- As a corollary, the sample concentration does not need to be accurately known in order to obtain a viable measurement. This can be of enormous help when dealing with in-process or formulated product samples.
- There are minimal sample requirements. Typical injected samples are 100–200 μL with a net concentration in the mg/mL range. This can be advantageous when a sample is limited (say in debugging processing details), but it can have a

negative side with respect to preparing a representative sample for analysis, as the material is very heterogeneous. Naturally, subsampling of a large parent solution can readily address this issue.

The hydrodynamic size of a chain in solution with fixed topology (ie, linear or branched) will scale with molecular weight, so the operational separation is effectively one of molecular weight and provides for estimation of the various distribution quantities. Since the common detection schemes are generally concentration sensitive, the *i*th weight fraction (w_i) is the experimentally determined quantity, thus M_w is based on the actual analytical measurement results. In contrast, M_n and M_z rely on derived calculations from the acquired analytical with heavy weighting at either the low- or high-molecular-weight wings of the distribution, which significantly lowers the precision of their determination. For a well-controlled and calibrated system, long-term variability of M_w determinations are generally better than $\sim$5% relative with M_n, and M_z often $\sim$10–15% precision. Note that these are rough results. Observed precision is composition dependent. Sample requirements, such as the need to employ a binary mobile phase to obtain a fully soluble form, can impact precision. Finally, shear forces in an analytical SEC column can be quite high. SEC system flow rates suitable for low- and moderate-molecular-weight materials (1.0 mL/minutes for typical $\sim$7–8 mm inner diameter SEC columns) may yield in situ shear degradation of the higher-molecular-weight grade materials available. The obvious approaches of either lowering the flow rate (to 0.25–0.50 mL/minutes) or using SEC packing materials with a larger particle size will remedy this problem.

By and large, routine SEC-based molecular weight determinations generally yield molecular weight averages that refer to chemically distinct narrow molecular weight distribution calibration standards used in calibrating the chromatographic system elution volume–molecular weight relationship. As such, reported molecular weights are often referred to as relative or equivalent molecular weights. Values provided by many laboratories adhere to this methodology. Finally, while most materials mentioned (aside from sodium carboxymethyl cellulose) are nominally nonionic, it is always strongly recommended that users employ some level of dissolved indifferent electrolyte in the SEC mobile phase. The reason for this is that even nominally nonionic polymers may possess low levels of carboxylate functionality due to adventitious oxidation events during preparation or aging. In the absence of screening indifferent electrolytes, chains that possess even a low level of charge may be electrostatically excluded from the internal pore volume of the packing material. In the absence of molecular weight-sensitive detection, this occurrence can yield incorrect molecular weights, since the elution of the sample is now influenced by factors outside of size alone.

More recently, SEC separations have been combined with molecular weight-sensitive detection schemes (online light scattering, online viscometry) to enable integrated development of operational absolute molecular weight–elution volume calibration of the system and online determination of absolute molecular weight information. The potential compositional complexity of the copolymeric systems used as pharmaceutical excipients renders light scattering determinations of molecular weight rather more complex than simple characterization of typical synthetic homopolymer systems. While the absolute nature of light scattering molecular weight determinations are often touted, that ideal is often not completely realized with the complex copolymeric systems discussed here. Furthermore, absolute molecular weight determinations are quite dependent upon the determination of an accurate, specific refractive index increment (dn/dc). Inspection of the literature for tabulated values of this quantity often shows a wide range of values in rather similar solvent systems. Ultimately, the accuracy of absolute molecular weights often relies on the quality of this secondary quantity. Under well-controlled conditions using purified samples, dn/dc can be determined with high precision. The usual secondary issue associated with this is how strongly dn/dc varies with composition and to what extent composition of the samples also varies.

While productive use of molecular weight-sensitive detection relies on judicious understanding of all the factors required to generate quality data, one overriding value of the approach is that if adsorption of the analyte occurs on the stationary phase packing material, the result is unaffected to the degree that elution zones may overlap with genuinely low-molecular-weight material. In routine SEC, which relies on the use of a secondary calibrant to allow creation of a MW/elution volume calibration curve, there is no direct indication if analyte is adsorbing to the stationary phase, and this will naturally have an enormous impact on any calculated results.

Although SEC is certainly one of the most used approaches to polymer molecular weight determination, the batch classical methods are still used.

Colligative property measurements include both vapor pressure and membrane osmometry. While both measurements rely on colligative effects, in a practical sense, they complement one another rather well.

Owing to the physical limitations of membranes, membrane osmometry is ill suited for relatively

low-molecular-weight polymers (M_n <25,000 Da), and finds its main utility with relatively high-molecular-weight materials and, under carefully controlled conditions, polyelectrolytes. Since low-molecular-weight contaminants are able to equilibrate across the membrane, membrane osmometry is insensitive to low-molecular-weight contaminants. It is also insensitive to chemical heterogeneity, which is a major advantage in the analysis of copolymeric materials.

In contrast, vapor pressure osmometry tends to excel in the lower-molecular-weight region (M_n <20,000 Da), for which membrane osmometry is ill suited. However, the technique is very sensitive to the presence of any low-molecular-weight impurities, which requires the use of highly purified samples and is inapplicable to polyelectrolyte systems.

Although colligative property measurements have a well-founded physical basis, any methodology that allows an experimentalist to count molecules in an assembly can be used to provide an estimate of M_n. Other techniques that have seen utility in the literature include end group analysis, in which unique functionality or reactivity or end groups are used as a mechanism to count the total population of chains in a sample.

At the other end of the spectrum, analytical ultracentrifugation can be used to quantify the M_z of a sample population. While this technique has seen substantial use in the analysis of proteins and other biologically derived polymers and has had a recent resurgence with the commercial rerelease of the Beckman Coulter Optima XL-A series of instruments, it remains something of a niche technology despite the appreciable technical benefits (lack of a potentially adsorptive stationary phase as in SEC, wide range of solution conditions possible, ability to assess complicated equilibria, and ability to handle ultrahigh molecular weight and colloidal systems under conditions of very low sample shear).

The measurements of M_w averages tend to dominate molecular weight characterization of polymers. For the sake of this discussion, we will consider methods that directly measure M_w together with methods such as viscometry, which yield a response that generally tracks M_w.

For the determination of an absolute M_w, light scattering, either alone or in conjunction with SEC, is the primary experimental approach. The use of SEC with either narrow distribution molecular weight standards or inline viscometric detection with universal calibration transformation, provides results that can range from a simple relative molecular weight, if the calibration standards are chemically distinct from the analyte, to near absolute M_w results.

Potential pitfalls of light scattering include the possible impact of dust or suspended precipitates for batch-style measurements, association phenomena that result in the inflation of the experimental results, the use of binary solvents to improve polymer solubility or the lack of strict adherence to constraints required to guarantee that absolute results are obtained (chemical homogeneity of the chain populations, maintaining constant chemical potential of counterions for polyelectrolyte samples (often approximated by the use of "swamping" electrolyte in large relative excess), accurate values of dn/dc).

Pairing light scattering with a prior SEC separation adds substantial power to both techniques. While SEC provides a means to separate material according to its hydrodynamic volume in solution, it also provides an excellent means to filter the analytical sample to eliminate dust and other artifacts from the light scattering signal.

While the underlying fundamental physical quantity of interest is the molecular weight of a material, and most performance traits are related to the molecular weight, most materials are classified in a more operational manner via grade classification according to a viscosity measurement under specified conditions. Naturally, viscometry or rheology can be used as a method to quantify polymer molecular weight, as well as to provide other insights into the nature of the material.

7.2.3 Rheological properties

Rheology is the study of the deformation and flow of a material when subjected to an applied force. The practical consequences of rheology play out in everyday experiences: in the kitchen as one thickens a water-based fluid with starch or forms a jelly through the use of gelatin dissolved at high temperature. In both extremes, the mechanical properties of the final preparation are manipulated through the judicious addition of a polymeric material. The science of rheology attempts to quantify the material changes exemplified by these systems. Detailed treatments of the complete scope of rheology are readily available.[17–20] Dealy has also prepared a convenient summary of official nomenclature and basic terms related to rheological science.[21]

In assessing the response of a material to an applied force, it is critical to appreciate that the applied force can span many orders of magnitude, ranging from simple gravity acting upon a particle suspended in a fluid medium to the very high shear rates experienced in high-speed mixing or high deformation rates experienced in high-speed conveyance and manufacture of solid products. Depending upon the particular details in manufacture and of usage, this entire range of shear

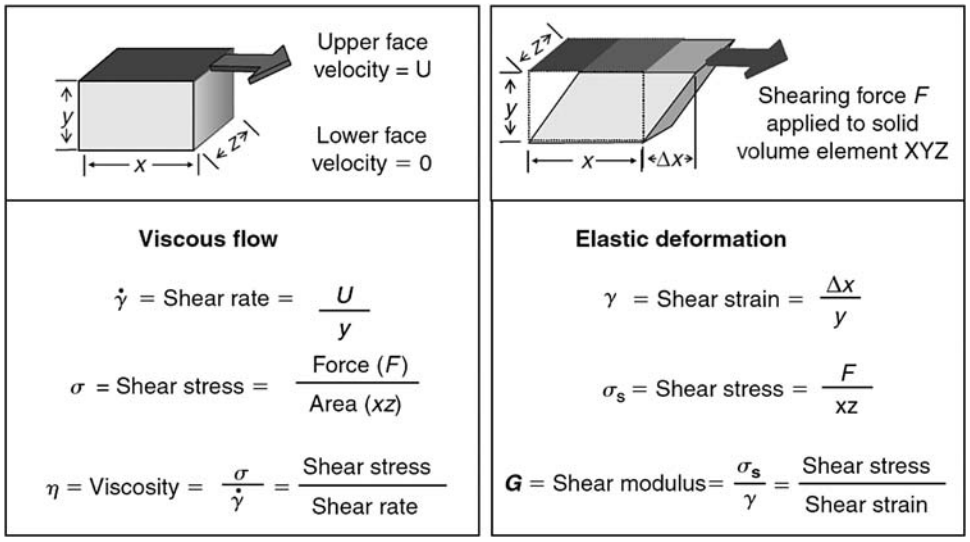


FIGURE 7.6 The basic mechanics of the flow of viscous liquids and deformation of elastic solids. In viscous flow, the energy input into the system is dissipated, while it is stored in the pure elastic deformation of a solid. Real materials are viscoelastic in nature, displaying both viscous and elastic character, with the balance determined by both the magnitude and duration of the applied force.

or deformation rates may be sampled at various points in routine usage.

Rheology considers not only the viscous flow of liquids but the elastic and plastic deformation of solids. In terms of the essential mechanics, the physical situations are quite similar (see Fig. 7.6). Real materials possess both elastic and viscous traits, with the key feature being the timescale over which the shear force is applied relative to the intrinsic relaxation time for the material.

In most cases, the response of a system is quantified by measuring the shear stress (σ, Pa) that develops in a system under the conditions of a controlled and specified shear rate ($\dot{\gamma}$, second^{-1}) for viscous flow or shear strain (γ) for pure elastic deformation. Typically, one speaks of the viscosity (η) of a fluid, which is simply the ratio of the shear stress divided by the shear rate:

$$\eta = \frac{\sigma}{\dot{\gamma}}$$

or the shear modulus (G):

$$G = \frac{\sigma}{\gamma}$$

in the case of solids.

If we restrict our attention to liquid systems for the moment, in the simplest case, which experimentally applies to simple, low-molecular-weight, nonassociating liquids, flow is directly proportional to the applied force. Equivalently, the ratio of the shear stress to applied the shear rate is a constant independent of the applied shear rate (see Fig. 7.7). Systems that behave in this fashion are termed Newtonian. Although it is not explicitly noted, Newtonian fluids will also display a time invariant viscosity. If the shear rate is maintained constant over time, the temperature is maintained constant, and chemical degradation of the sample does not occur, the shear stress that results, and hence the measured viscosity, is also constant. Although most common Newtonian liquids are also low viscosity, these two traits are not inexorably tied together. For example, concentrated sucrose solutions possess substantial viscosity and also exhibit Newtonian flow.

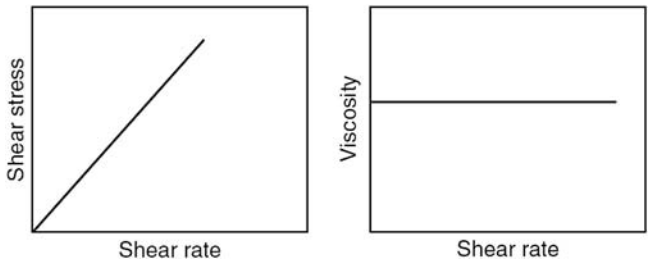


FIGURE 7.7 The simplest case for fluid flow is provided by a simple Newtonian liquid. In this case, the fluid viscosity is independent of the applied shear rate. Common low-molecular-weight unassociated liquids and solutions are Newtonian.

When a polymer is dissolved in a low-molecular-weight Newtonian solvent, the rheological response can become decidedly more complex. The most significant effect is that one observes that the solution viscosity is no longer independent of the applied shear rate under all conditions. One will generally observe that at vanishingly small shear rates, there will be a limited low-shear Newtonian plateau over which the solution viscosity appears independent of the applied

shear rate. As the applied shear rate is increased, one will enter a shear-thinning regime in which the apparent viscosity drops as the shear rate is increased (see Fig. 7.8). Finally, at very high shear rates, one may observe the reestablishment of a regime in which the apparent viscosity is again independent of the applied shear rate, yielding a limited high-shear Newtonian plateau.

An example with actual data is provided in Fig. 7.9. This profile shows the shear-thinning behavior of various viscosity grades of 2.0 wt.% hypromellose in water. Note that even at rather low effective shear rates, achievement of a well-defined low-shear Newtonian plateau is not assured. Finally, the fairly strong shear rate dependence

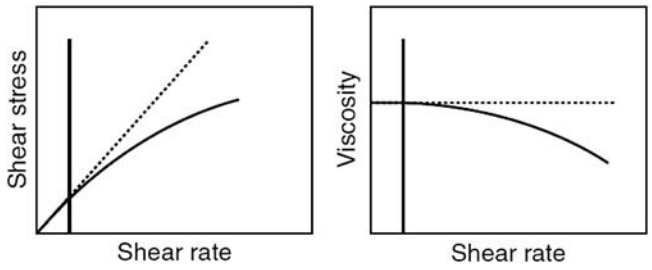


FIGURE 7.8 Most solutions of polymers exhibit some degree of non-Newtonian flow behavior. Simple shear thinning is the most straightforward case. In this example, the system is Newtonian until a critical shear rate is achieved. Above this critical shear rate, the apparent viscosity drops with further increases in shear rate.

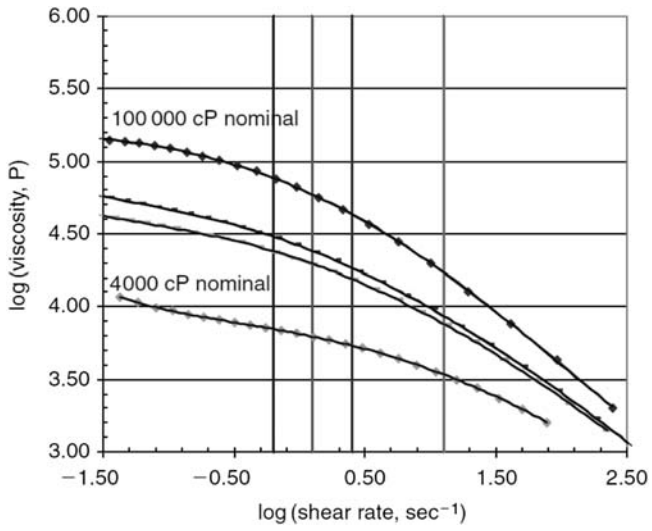


FIGURE 7.9 Typical shear-thinning behavior exhibited by 2 wt.% hypromellose solution in water. The minor apparent upturn in the apparent viscosity of the 4000 cP nominal viscosity sample is an instrumental artifact due to a low analytical signal. The vertical lines indicate (left to right) the effective shear rate applied in a Brookfield rotational viscometer equipped with a cylindrical fixture at 3, 6, 12, and 60 revolutions per minute (rpm).

in the measured viscosity of a polymeric solution can have very practical consequences. As can be observed from the estimated shear rates for a determination of viscosity using a simple Brookfield rotational viscometer under rotational speeds commonly employed in laboratory work (3, 6, 12, and 60 rpm), for a shear-thinning fluid, precise adherence to specified conditions is a clear requirement if the direct comparison of results is desired. For example, the apparent viscosity of the 100,000 cP (grade designation based on 2% Ubbelohde viscosity using the USP methodology) would appear to be roughly 75,000 cP if measured at with a cylindrical spindle at 3 rpm, while the apparent viscosity is less than 20,000 cP if measured at 60 rpm. This is the reason that the shear rate of the measurement needs to be known in assessing viscometric results, particularly when the sample is in the semidilute regime. Furthermore, while factors such as spindle rotation rates obviously impact the shear rate applied to a system, less obvious factors such as vessel size (if small relative to the rotational fixture in a Brookfield viscosity measurement) can also contribute to the actual shear rate applied to the sample.

In addition to the pronounced shear rate dependence in the apparent viscosity, the concentration dependence can be broken into two regimes for nonelectrolyte polymers, with some added complexities apparent in the case of polyelectrolytes. If one views the state of a discrete polymer chain as the concentration of a system is increased, the solvated chain will rotationally sweep out a distinct and large volume element of the solvent. In a dilute solution, on average, all of the chains will be able to rotationally average their motions without any hindrance from neighboring polymer chains. However, as the concentration is increased, one will reach a point at which the rotationally averaged volume elements first touch, then overlap. Once these volume elements start to overlap, the chains start to entangle with one another. When this occurs, the concentration dependence of the apparent viscosity at constant shear rate increases substantially. This regime, under which polymer chains overlap and entangle with one another, is referred to as the semidilute concentration regime. If one were to examine the profile of apparent viscosity against concentration in log–log coordinates, two distinct power law regimes are generally apparent in the dilute and semidilute regimes. Power law exponents are generally close to 1 in the dilute regime and increase to 3–4 as one transitions to the semidilute regime (see Fig. 7.10). From a fundamental basis, characterization of the apparent viscosity dilute and semidilute regimes is best done using results extrapolated to zero shear rate (ie, reflective of the low-shear Newtonian plateau). However, operationally, these analyses and general guides also apply with reasonable fidelity under

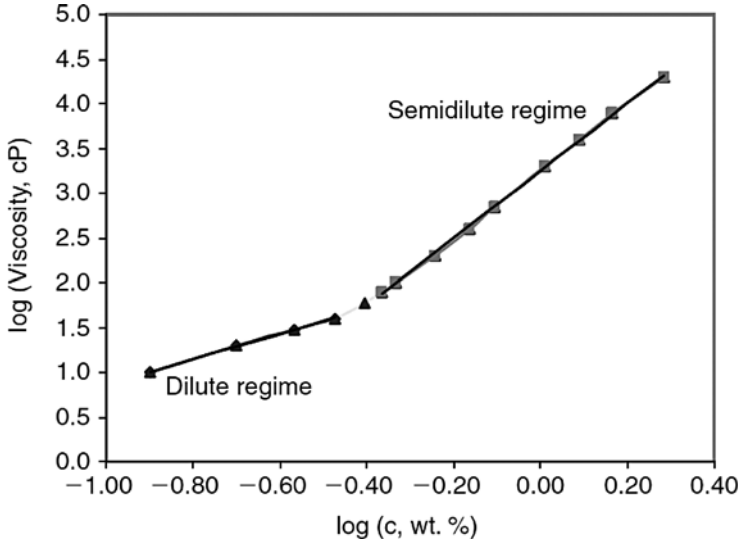


FIGURE 7.10 Example of the concentration dependence of viscosity for sodium carboxymethyl cellulose (DS ~ 0.7, 7H type) as concentration is varied from roughly 0.10 to 2.0 wt.%. Clear distinction of the dilute and semidilute regimes is apparent. The solutions represented in this example were prepared using pure water as the solvent, hence the ionic strength of the solution is not held constant as the polymer concentration is increased. This results in a slight lowering of the power law slope in both the dilute and semidilute regimes.

conditions of finite shear. The distinction is important, since most workers will have access to a Brookfield rotational viscometer to characterize these materials, with this style of measurement generally being practiced at effective shear rates well above that needed to observe the low-shear Newtonian plateau.

In the case of polyelectrolyte, additional complications exist due to the strong dependence of the expansion of a solvated chain on the net ionic strength of a solution. Fig. 7.11 depicts the dependence of the intrinsic viscosity of a sodium carboxymethyl cellulose sample as a function of ionic strength in dilute solution. As one would anticipate, the addition of an indifferent electrolyte screens the repulsion of like charges along a polyelectrolyte backbone, allowing the chain to adopt a more relaxed random coil configuration in solution. The net result is a drop in viscosity in the system. As depicted in Fig. 7.11, the drop in viscosity can be relatively large for modest changes in ionic strength.

In a solution with no added indifferent electrolyte, the general viscosity behavior of polyelectrolytes is similar to nonelectrolytes, with the added feature of the progressively increasing screening of bound charge sites yielding an apparent lowering of power law exponents in both the dilute and semidilute regimes, since the apparent ionic strength increases with polymer concentration. This complexity can be suppressed by the addition of an indifferent electrolyte to maintain a nearly constant net ionic strength in solution. Naturally, this can further suppress viscosity build in the system.

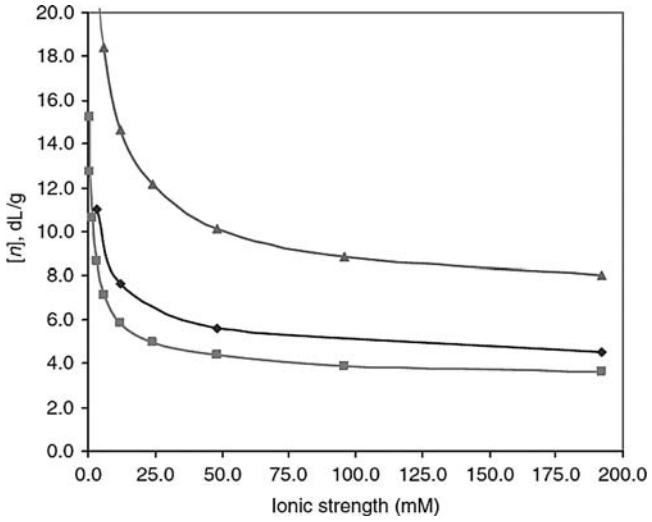


FIGURE 7.11 Impact of an added electrolyte on the intrinsic viscosity of a typical polyelectrolyte, in this case sodium carboxymethyl cellulose (NaCMC). (Original data from Pals, D. T. & Hermans, J. J. (1952). Rec. Trav. Chim. Pays-Bas, 71, 433–443.) At low ionic strength, the NaCMC chains adopt a relatively rigid, rod-like conformation due to electrostatic repulsion between adjacent chain-bound anionic charges. As the ionic strength is increased, this repulsive interaction is screened, and the polymer adopts a less rigid conformation.

Finally, in addition to solution shear thinning due to alignment of chains by the applied shear field, the presence of labile association aggregates can yield a time-dependent apparent viscosity that is dependent on the recent shear history of the sample (see Fig. 7.12). This characteristic is termed "thixotropy," and it arises

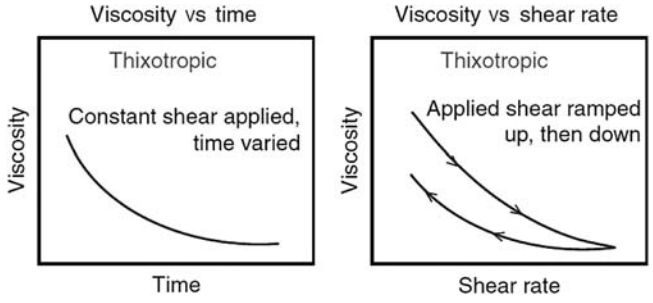


FIGURE 7.12 Polymeric systems in which reversible interpolymer association is possible display a characteristic known as thixotropy. Basically, the mechanical stress applied to the system causes reversible disassembly of the aggregate network. Under conditions of steady shear, this yields a slow decrease in the apparent viscosity over time. The extent of disassembly is dependent on the net force applied to the system, hence the drop in viscosity can occur from either a quiescent state or if a sheared system is subjected to a higher shear rate. By the same token, on a decrease in applied shear rate, the apparent viscosity exhibits recovery. The magnitude of the effect is generally quantified by measuring the disparity in flow curves obtained in a simple experiment in which the shear rate is ramped up then down with continuous recording of the viscosity.

from the association-dissociation of aggregates, which is much slower than the timescale of the measuring experiment.

Although relating solution rheology to molecular weight is best pursued using extrapolated zero shear values, the use of power law relationships in the semidilute concentration regime at finite shear typically yields operationally useful relationships for the quick estimation of molecular weight. Owing to pragmatic instrumental limitations, these experiments generally involve data acquired over a range of shear rates. This can lower the overall quality of the correlation, but it is still useful for routine estimates.

Power law relationships of the form:

$$\log(M_w) = a + b * \log\left[\frac{\eta}{c^{\alpha}}\right]$$

are often found to operationally model the data. In the equation above, M_w is the average molecular weight, while η is the solution viscosity at concentration c. The constants a, b, and α can be viewed simply as mathematical fitting constants, although α is related to the power law concentration exponent in the semidilute regime. An example of this treatment is shown in Fig. 7.13 for three types of cellulosic derivative. The specific relationships shown are as follows:

HPC:

$$\log(M_w) = 5.08 + 0.305 * \log\left[\frac{\eta}{c^{3.2}}\right]$$

Hydroxyethyl cellulose:

$$\log(M_w) = 5.00 + 0.306 * \log\left[\frac{\eta}{c^{3.1}}\right]$$

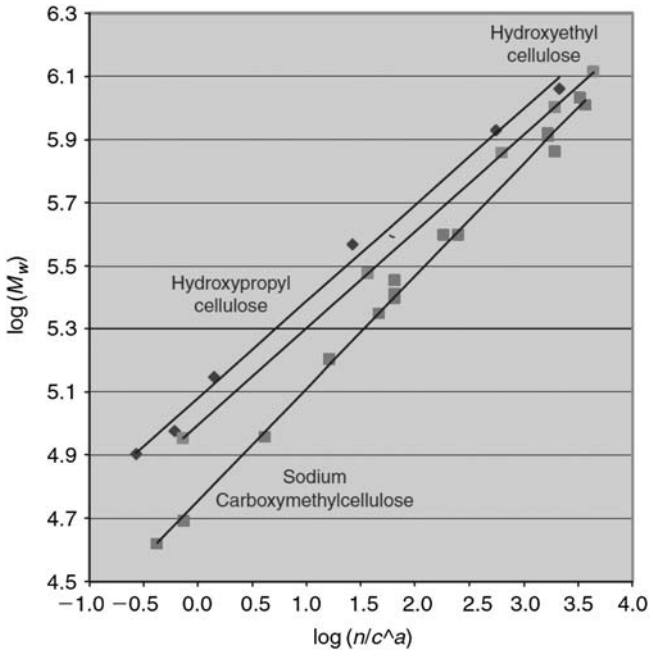


FIGURE 7.13 Polymer master viscosity/concentration/M_w profiles for cellulose ether excipients using Brookfield viscosity values at variable concentration and assuming a simple power law relationship. Separate curves for HPC, hydroxyethyl cellulose, and sodium carboxymethyl cellulose are shown. These curves span the entire product range viscosity for the derivative types shown and listed in Table 7.8, later in this chapter.

Sodium carboxymethyl cellulose:

$$\log(M_w) = 4.78 + 0.356 * \log\left[\frac{\eta}{c^{3.1}}\right]$$

where:

M_w is the nominal material molecular weight;
η is the solution Brookfield viscosity;
c is weight percentage polymer concentration.

Naturally, these relations can be rearranged to yield initial estimates of the concentration of any specific grade of polymer required to achieve a target of viscosity.

There are a few key points related to rheology and polymers that are worth noting:

1. Since the rheology of a polymer solution is molecular-weight dependent, rheology/viscometry can be used as an indirect proxy to monitor molecular weight.
2. Although, strictly speaking, one should use the apparent viscosity extrapolated to zero shear in modeling the molecular-weight dependence of polymer solution viscosity, reasonably good master curves obtained at finite shear rates and at variable concentration can be used to provide decent pragmatic estimates of the molecular weight of a polymer sample.

3. Owing to the very strong concentration dependence, the actual polymer concentration needs to be accurately known. Since the power law exponent of viscosity–concentration curves in the semidilute region have a log–log slope of 3–4, an x% error in specifying the concentration yields a 3x%–4x% error in the calculated viscosity.
4. Solution viscometric experiments generally provide information related to the average M_w of a system. No information on the breadth of the molecular weight distribution is provided.
5. The magnitude of the exponent in the power law dependence in viscosity as a function of concentration, as seen in Figs. 7.10 and 7.13, is the origin of typical rules of thumb used to estimate viscosity at different concentrations. For example, a doubling in weight percentage concentration often yields an increase in viscosity of roughly 10-fold, which follows directly from the type of power law behavior shown in Fig. 7.10.

7.2.4 Polymers in solution

Understanding the unique solution behavior of polymers mainly resides primarily in appreciating a few key points.

First, the high molecular weight of a polymeric solute significantly lowers the entropy of mixing for a polymer solvent binary system, lowering this component of the net free energy of solution. This was treated some time ago by Flory[6] and more recently by Sanchez and Lacombe,[22–24] which captures the essential results.

At an ideal level, taking a solution as a simple lattice model, we find:

$$\Delta S_{\text{mix}} = -R(n_1 \ln(\varphi_1) + n_2 \ln(\varphi_2))$$

or

$$\Delta S_{\text{mix}} = -R\left[\frac{\varphi_1}{DP_1}\ln(\varphi_1) + \frac{\varphi_2}{DP_2}\ln(\varphi_2)\right]$$

As is apparent from the inverse dependence of the prelogarithmic factor of each term on the degree of polymerization (DP) or, equivalently, the molecular weight of the components in the mixture, if either one or both of these components is a high-molecular-weight material, the magnitude of the entropy is lowered as the molecular weight is increased. In broad strokes, as the molecular weight increases, the free energy driving force for dissolution is lowered, since the gain in entropy by the polymer on dissolution is quite low.

One direct consequence of the lowered entropy of solution is that in a chemically homogeneous system precipitation of a polymeric solute, as solvent conditions transition from good to poor, it starts with the high-molecular-weight portion of the distribution and proceeds with decreasing molecular weight. This provides a relatively facile manner to perform crude molecular weight-based preparative isolation. It also suggests that for a partially soluble material (again, assumed to be chemically homogeneous), the undissolved fraction is typically enriched in the high-molecular-weight fraction.

Secondly, most polymers under consideration here are copolymeric materials, for which the various component monomers have rather different intrinsic solubility. For example, in the case of HPC, a single unsubstituted anhydroglucose moiety is quite hydrophilic, while a heavily hydroxypropylated anhydroglucose residue is rather hydrophobic. The constituent monomers in a real sample will span a wide range in hydrophobic nature, thus their distinct distribution across chains (interchain heterogeneity), and within chains (intrachain heterogeneity). The net result is that at any specific solvent composition a situation can arise in which subpopulations of entire chains may be soluble or insoluble, or discrete subsections of single chains may be solvated or unsolvated.

Thirdly, the linear array of monomer groups provides a spatial template upon which relatively low-energy intermolecular forces can operate in a cooperative fashion to yield fairly strong interchain associations. A specific example of this would be the interchain association of unsubstituted anhydroglucose runs in sodium carboxymethyl cellulose to yield textured solutions and thixotropic rheology.

With respect to gross solution behavior, the same basic rules that guide the dissolution of low-molecular-weight species also apply to polymers. As with low-molecular-weight materials, “like dissolves like.” This is captured somewhat more quantitatively in the extended solubility parameter approach,[25–27] in which the cohesive energy of a material is divided into contributions arising from distinct types of intermolecular interactions (dispersive, polar, hydrogen bonding). While the usual admonishment that “like dissolves like” is somewhat more quantitatively expressed within the solubility parameter approach as a matching of the cohesive energy densities of the two materials (solubility is maximized when the cohesive energy densities of the solvent and solute are equal and decreases as they become increasingly disparate), the categorization of the extended solubility parameter approach places a rationale around the basic picture in terms of discrete interaction mechanisms that exist between molecules.

Unlike the situation with low-molecular-weight liquids, one cannot experimentally determine the molar

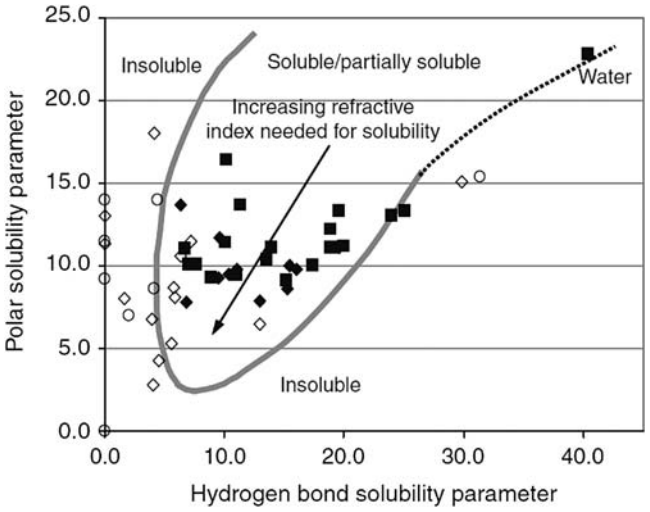


FIGURE 7.14 Solubility characteristics of hydroxypropyl cellulose as a function of polar and hydrogen bond solubility parameters. In this solubility map, the solid curve is the soluble/insoluble phase boundary presented in reference.[28] The dashed line is an interpolated boundary taking into account water solubility characteristics. Points shown represent the specific solvents listed in references[27–29] as yielding solutions, described as clear and smooth/soluble (■), moderately granular and/or hazy/partially soluble (◆) or insoluble (○), or uncharacterized systems of unknown solubility (◇).

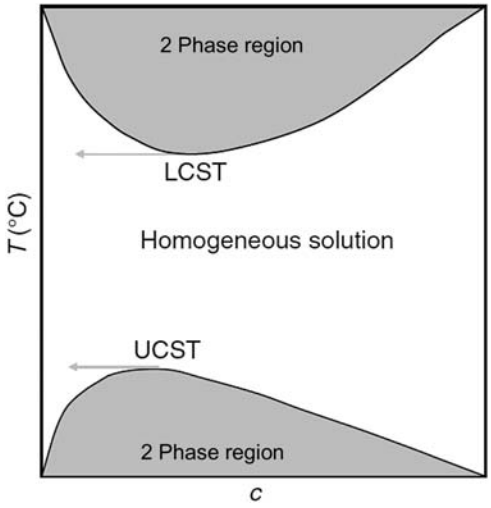


FIGURE 7.15 Schematic representation of upper and lower critical solution temperatures, UCST and LCST respectively. A UCST exists when a polymer transitions from partially soluble to soluble as temperature is raised, while an LCST occurs when a single homogeneous phase splits into two phases with the same increase in temperature. An LCST is often seen with relatively hydrophobic materials, while UCST behavior tends to be observed with systems involved in specific self-association phenomena.

energy of vaporization of polymers and using the molar volume calculate a direct cohesive energy density ($\delta = (E_{vap}/V)^{0.5}$). However, one can perform challenge–response experiments to assess the solubility behavior of a polymer in a number of low molecular weight liquids (insoluble, partially soluble/swellable, soluble) and use those observations to construct solubility parameter-based maps, which can aid in the selection of optimum solvent systems for dissolving a polymeric material. As a specific example, consider HPC.[28,29] Fig. 7.14 contains the qualitative solution behavior of HPC in a variety of solvents presented as a two-dimensional solubility map in terms of the polar and hydrogen bond solubility parameters. As is apparent, there is a reasonably clear demarcation between the insoluble and partially or completely soluble regions. Note that the partially soluble and soluble regions exhibit a substantial overlap. This is largely a consequence of the two-dimensional nature of the plot, which neglects the explicit contribution of dispersive interactions to solubility. The specific solvents for which HPC exhibits good solubility at low polar and/or hydrogen bond solubility parameters tend to be high refractive index liquids (aromatic or halogenated), which increases the dispersive contribution of the net solution energy.

The dissolution characteristics of the polymeric species become complicated when the constituent monomers have vastly different solvation requirements. In that case, the usual recommendation of "like dissolves like" may only apply to a portion of a given polymer chain. In this case, the use of binary and/or ternary solvent mixtures can often yield solutions not possible from any of the pure solvent components.

One other aspect of polymer solubility is the existence of upper and lower critical solution temperatures, referred to as UCST and LCST, respectively, seen in aqueous solutions of many of these polymers. The situation is diagrammatically shown in Fig. 7.15. In the case of an LCST, the system splits into a two-phase system comprised of polymer-rich and polymer-lean compositions at high temperature. As the temperature is lowered, the composition of these two phases approaches one another and then merges into a single phase at the LCST. Similar behavior occurs with the UCST, except that the transition from a two-phase to a one-phase system occurs as the temperature is raised.

In the current application settings, one typically observes an LCST from the perspective of a cloud point, which is simply due to performing the experiment in reverse—starting with a homogeneous solution in the single-phase region and raising the temperature until the solution splits into two phases. This occurs uniformly throughout the volume of solution, which typically yields small droplets of the polymer-rich phase dispersed in the polymer-lean phase. This generates a high level of scattered light, rendering the fluid opaque, and is the origin of the term "cloud point."

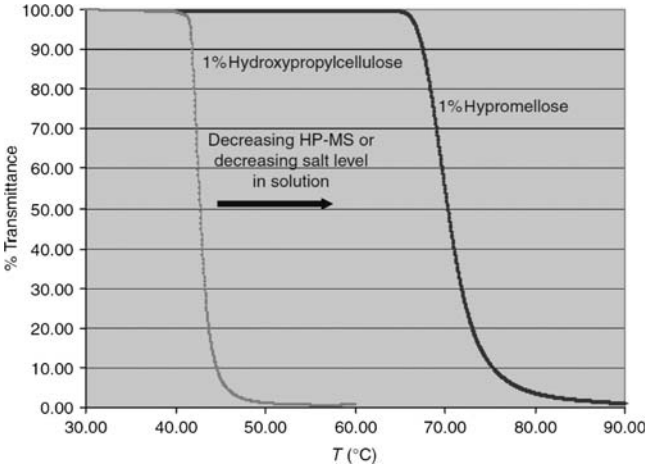


FIGURE 7.16 Cloud point curves for typical high-viscosity grades of hydroxypropyl cellulose and hypromellose (hydroxypropylmethyl cellulose). The major difference in cloud points between these two materials is reflective of the differing hydrophobicities (determined by a combination of hydroxypropyl and methoxyl groups for hypromellose). In addition to the native polymer composition, addition of an indifferent electrolyte can lower the cloud point by a few degrees for typical physiologically relevant ionic strength.

Fig. 7.16 provides an example of two chemically similar systems, HPC and hypromellose, both as 1% solutions in water. The substantial difference in cloud points is related to the substantially different hydrophobicity of these two materials. As is typical for marginally soluble hydrophobic materials, the addition of an electrolyte will move the cloud point to a lower temperature, as would an increase in the hydroxypropyl content of either system or the methoxyl level of the hypromellose sample.

7.2.5 Polymer morphology and physical properties

Polymers may consist of both crystalline and amorphous regions; the degree of crystallinity is usually expressed in terms of a weight fraction or volume fraction of the crystalline portion. The crystallinity of polymers is characterized by their degree of crystallinity and, in principle, can range from zero for a completely noncrystalline or amorphous polymer to one for a theoretical completely crystalline polymer; however, it typically ranges from 10% to 80% in real polymers.[30] Therefore, such crystalline polymers are often called semicrystalline. The properties of semicrystalline polymers are determined by the degree of crystallinity and by the size and orientation of the polymer chains. Polymers containing microcrystalline regions (ie, semicrystalline polymers) are generally tougher (can be bent more without breaking) and more impact-resistant than totally amorphous polymers.

The term melting point, when applied to crystalline polymers, suggests a transition from a crystalline or semicrystalline phase to an amorphous melt phase. Although abbreviated as simply T_m, it may be more properly called the crystalline melting temperature. Among synthetic polymers, crystalline melting is only associated with linear thermoplastic polymers, as highly cross-linked thermosetting polymers will decompose at high temperatures rather than melt.

In an amorphous state, the arrangement of the polymer chains is completely random. There is no restriction to chain movement from the lattice ordering of partially crystalline polymers. Therefore, amorphous polymers do not have sharp melt points; instead they soften gradually with increasing temperature. The physical properties of amorphous polymers are related to the extent of their molecular mobility in the material, which is governed by the chain flexibility and the temperature of the system.

On cooling, rubbery materials undergo a liquid-glass transition, which has also been called a rubber-glass transition. The temperature at which this transition occurs is termed the glass transition temperature (T_g), which describes the temperature at which amorphous polymers undergo a second-order phase transition from a rubbery, viscous amorphous solid to a brittle, glassy amorphous solid. Drastic changes in the physical properties, such as hardness and elasticity, can be observed during this transition as it reflects the onset of changes in the cooperative thermal motion of the polymer backbone segments. These changes are completely reversible, since the transition is a function of molecular mobility rather than polymer structure. At temperatures above T_g, the polymer chains are in relatively rapid thermal motion. The chain movement becomes progressively slower as the temperature is lowered. At T_g, the polymer chains are essentially locked in the conformation they possessed before reaching T_g. Below T_g, the polymer is virtually frozen in a glassy state with a completely random structure. The T_g of a polymer depends largely on the chemical structure of the polymer chain. The glass transition temperature may be modified by altering the chain flexibility, monomer structure and copolymer composition, degree of branching or cross-linking in the polymer, or by the addition of plasticizers. Long-chain branches may increase polymer strength, toughness, and the T_g due to an increase in the number of entanglements per chain. Increasing chain length tends to decrease the chain mobility and increase the glass transition temperature (T_g). This is a result of an increase in polymer chain interactions such as van der Waals attractions and entanglements associated with increased chain length. These interactions tend to immobilize individual chains more strongly in situ,

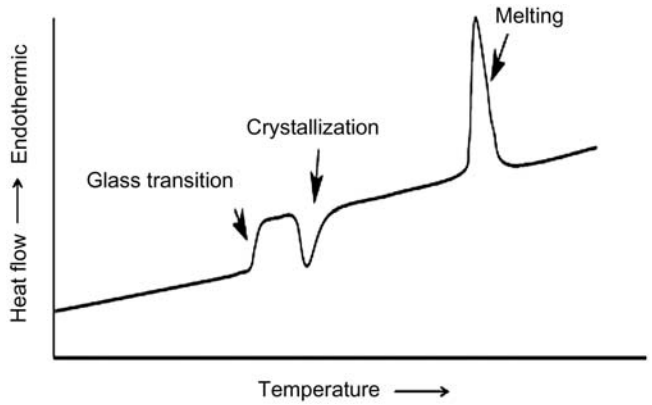


FIGURE 7.17 Typical differential scanning calorimetry (DSC) scan.

resulting in better resistance to deformation and matrix breakup, both at higher stresses and higher temperatures.

The glass temperature of a polymer can be characterized experimentally by measuring basic thermodynamic, physical, mechanical, or electrical properties as a function of temperature. Thermal method such as differential scanning calorimetry (DSC) is used routinely for this purpose. A typical DSC scan is shown in Fig. 7.17 where the heat flow is plotted against the average temperature. In addition to sharp peaks corresponding to crystallization and melting events, the change of slope of this scan occurs at the glass transition temperature, which can be related to the enthalpy change associated with this transition.

7.2.6 Structure–property relationships

In polymer chemistry and materials engineering, it is common to design novel materials and formulated systems with improved functional properties based on detailed study and understanding of how the structure of a material impacts the physical and chemical properties that are critical to the end use. In the pharmaceutical industry, it is common to apply this type of approach to the design of drug molecules. For example, quantitative structure–activity relationships are studied to enable the design of drug molecules with optimal therapeutic efficacy. A further example is the optimization of drug solubility and stability through the selection of appropriate salt and crystal forms. Similarly, a good appreciation of structure–property and structure-function relationships is also essential if formulators are to make rational, science-based choices when selecting polymers and other excipients as formulation components. Structure–property and structure–function relationships can be defined as understanding how the chemical or molecular structure of a polymeric excipient impacts the physical and chemical properties that are critical to the end use functionality. These concepts are gaining increasing significance in the context of quality by design (QbD). QbD is defined as a systematic, scientific, risk-based, and holistic approach to pharmaceutical development that begins with predefined objectives and emphasizes product and process understanding and process control. A requirement of QbD is to identify critical material attributes (CMAs), which are defined as the physical, chemical, and biological properties of an input material that should be within specific limits to ensure desired quality. For polymeric excipients, this essentially requires the understanding of structure-function and structure-property relationships.

The following section will attempt to illustrate some of the major, generally applicable structure–property relationships that are of practical use and specifically relevant to pharmaceutical polymers and their use in solid oral controlled-release dosage forms. These include the impact of MW, substitution, and copolymerization on solution, gel, mechanical, and thermal properties. While the general principles are applicable across a wide range of polymers, many of the examples given here will focus on cellulose derivatives and vinylpyrrolidone polymers, as these are the most commonly used classes of polymers in oral solid dosage form design.

7.2.6.1 Molecular weight effects

7.2.6.1.1 Effect of molecular weight on solution viscosity

Many polymer properties are molecular-weight dependent, including solution and melt viscosity, glass transition temperatures, mechanical properties, and gel strength. Perhaps the most important and frequently applied structure–property relationship is the correlation between dilute solution viscosity and molecular weight. As a longer polymer chain gyrates in solution, it can be visualized to occupy a larger volume of the solvent, collide, and overlap more frequently with other long polymer chains as compared to shorter polymer chains. The initial fundamental observation and prediction that solution viscosity is proportional to polymer molecular weight was made by Staudinger and Heuer.[31] For linear polymers, the Mark–Houwink equation provides an empirical model for this is relationship:

$$\eta = kM^a$$

where:

η is the specific viscosity;
M is the weight-average molecular weight;
K and a are polymer, solvent, and temperature dependent.

Generally, a varies between 0.5 and 1. For poor solvents in which the polymer remains coiled, $a = 0.5$. For good solvents, a varies between 0.65 and 0.8. For stiff, asymmetrical molecules in good solvents, a may approach 1. It should be noted this specifically applies to linear polymer molecules. Highly branched, bush-shaped polymer structures may have a large molecular weight but occupy proportionally smaller volumes in their solvated state.

7.2.6.1.2 Effect of molecular weight on mechanical and thermoplastic properties

Many physical properties of amorphous polymers, including mechanical and thermoplastic properties, show strong molecular-weight dependence. Frequently, a general relationship applies where molecular weight changes in the lower middle molecular weight range result in large changes in the physical property of interest. However, as shown in Fig. 7.18, after reaching a threshold level in molecular weight, further molecular weight increases result in only minor changes in the physical property of interest.

Often these phenomena can be explained on the basis of chain end concentration, that is, lower-molecular-weight polymer chains are shorter, and thus for the same amount of material, there is a proportionate increase in the number of molecules and thus chain ends. The chain ends tend to be associated with a greater degree of mobility and free volume as compared to the middle segments of a polymer chain. Consequently, lower molecular weight is generally associated with greater plasticity and lower glass transition temperatures. In contrast, longer polymer chains (lower chain end concentration) are generally associated with increased elasticity and flexibility.

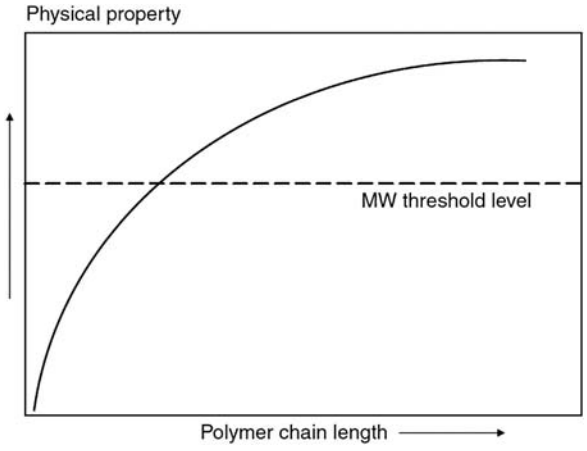


FIGURE 7.18 General relationship between polymer chain length (molecular weight) and physical properties such as tensile strength and T_g.

7.2.6.1.3 Mechanical strength of films

One of the most useful mechanical tests for polymeric and pharmaceutical materials in general is to determine its tensile strength and the accompanying stress–strain curve. This is generally done by using a mechanical testing machine and continuously measuring the force developed as the material is elongated at a constant rate of extension. Fig. 7.19 illustrates a typical stress–strain curve. Important mechanical properties of a material include the modulus (slope of the initial curve), which is a measure of the material stiffness; yield stress; strength; ultimate strength; elongation at break; and toughness (area under the curve).

For practical purposes, it is often useful to divide polymeric materials into five common classes depending on their stress–strain behavior. These common classes are illustrated in Table 7.1.

For polymer films, an increase in molecular weight tends to increase film tensile strength, elongation, and flexibility. This can be explained on the basis that longer polymer chains exhibit greater flexibility and elasticity. They can thus be extended further before rupture, as compared with short polymer chains.

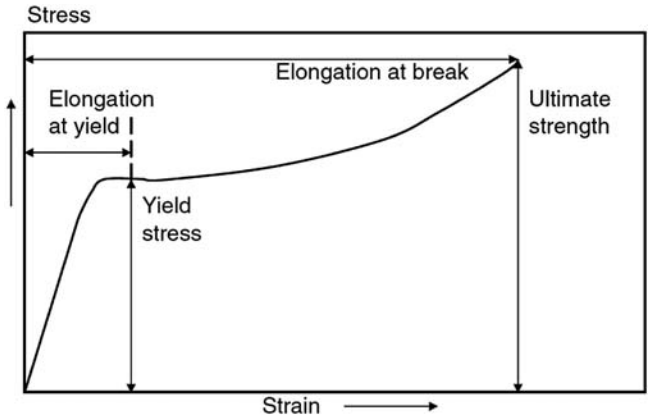


FIGURE 7.19 Typical tensile stress–strain curve for a polymer specimen tested in tension.

TABLE 7.1 Common Classes of Materials Based on Their Stress–Strain Curve Characteristics (Ref. 32)

Class	Modulus	Yield stress	Ultimate strength	Elongation at break
Soft and weak	Low	Low	Low	Moderate
Soft and tough	Low	Low	Moderate	High
Hard and brittle	High	None	Moderate	Low
Hard and strong	High	High	High	Moderate
Hard and tough	High	High	High	High

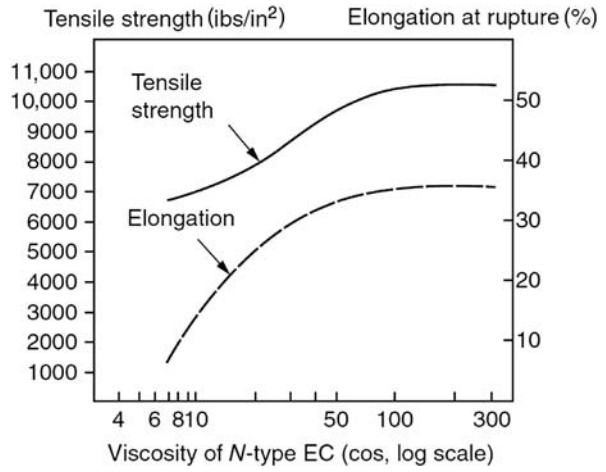


FIGURE 7.20 Ethyl cellulose film properties as a function of solution viscosity (molecular weight). All solutions viscosity values reflect 5 wt.% ethyl cellulose in an 80:20 blend of toluene:ethanol. *Courtesy of Ashland Specialty Ingredients Wilmington, DE.*

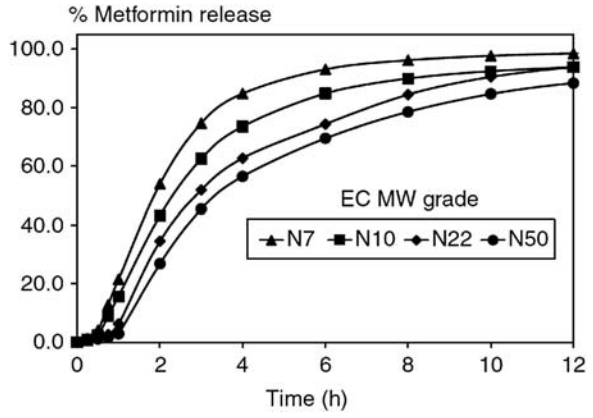


FIGURE 7.21 Drug release from metformin HCl pellets coated with various molecular weight grades of Ashland N Pharm ethyl cellulose (EC).[33] N7, N10, N22, and N50 refers to different molecular weight grades with 7, 10, 22, and 50 cps solution viscosity (5 wt.% solution in 80:20 toluene:ethanol), respectively.

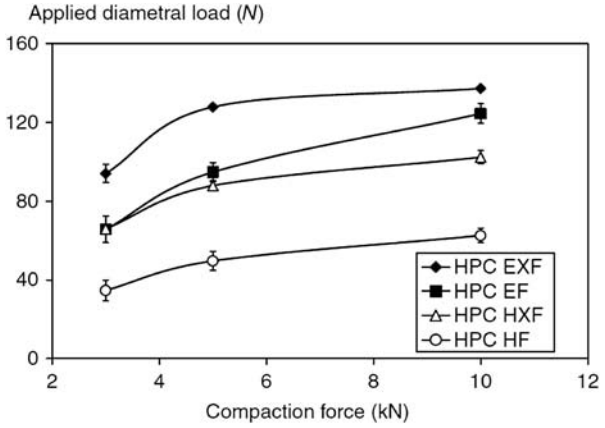


FIGURE 7.22 Compactability for various molecular weight and particle size types of hydroxypropyl cellulose. Klucel™ EXF Pharm HPC (80 kDA, 70 μm average particle size) Klucel™ EF Pharm HPC (80 kDA molecular weight, 200 μm average particle size), Klucel™ HXF Pharm HPC (1000 kDA molecular weight, 70 μm average particle size), Klucel™ HF Pharm HPC (80 kDA, 200 μm average particle size); 300 mg pure polymer tablets were compressed on a Manesty BetaPress using 7/16″ round flat-faced tooling.

A general relationship describing this behavior is given by the following equation:

$$\text{Tensile strength} = a - \frac{b}{M_n}$$

where a is the tensile strength threshold, beyond which further molecular weight increases become ineffective; b is a material-dependent constant; and M_n is number average molecular weight.[32] An example of this is given in Fig. 7.20, which shows the viscosity (MW) dependence of ethyl cellulose film properties.

Film strength is of significance in membrane reservoir controlled-release dosage forms such as multiparticulates or tablets coated with ethyl cellulose controlled-release films. It has been shown that the drug-release kinetics of such systems are directly proportional to the film strength and, by inference, polymer molecular weight (Fig. 7.21). This is due to the fact that the drug release from membrane–reservoir systems is osmotic pressure dependent. As the water enters the dosage forms via osmosis and builds up hydrostatic pressure in the core, it increasingly exerts a stress on the membrane until fissures and cracks open up, resulting in the release of the drug due to osmotic pumping.[33–35]

Similar examples of increasing film strength with molecular weight are known for the majority of pharmaceutical polymers. A further example is the report by Prodduturi et al., describing the increased mechanical strength of drug-loaded buccal films with increasing molecular weight of HPC used.[36]

7.2.6.1.4 Mechanical strength of tablets

As previously indicated, longer polymer chains impart greater elasticity, while the greater chain-end concentration found with decreasing molecular weight results in greater plasticity and molecular mobility. For directly compressed tablets, lower polymer molecular weight, therefore, results in less postejection elastic recovery and greater permanent deformation. For example, low-molecular-weight HPC grades have been reported to yield stronger tablets than high-molecular-weight HPC due to greater plasticity and lower postcompaction ejection.[37] However, this disadvantage of high-molecular-weight polymers can be partly neutralized by using fine-particle-size HPC (Fig. 7.22).[38] Similar observations have also been made for ethyl cellulose matrix tablets.[39]

7.2.6.1.5 Glass transition temperature, melting point, and melt index

The glass transition temperature (T_g), melting point temperature, and melt index are fundamental characteristics related to thermoplastic behavior. Similar to mechanical properties, a negative inverse relationship has also been postulated for T_g and molecular weight[40]:

$$T_g = T_g^{\infty} - \frac{k}{M_n}$$

where T_g^{∞} is the glass transition at infinite molecular weight, and k is a constant. The greater flexibility and lower T_g of lower-molecular-weight polymers are attributed to the greater concentration and mobility of polymer chain ends versus the chain middle. Due to their greater mobility and the effective defect in material packing that they create, chain ends are associated with greater excess free volume. In a series of equally substituted polymers, lower-molecular-weight polymers will thus have a greater concentration of chain ends, resulting in greater excess free volume, greater molecular mobility, and a lower T_g. A good example of this is provided by poly(vinyl pyrrolidone) (PVP), a simple, linear, polymer molecule with additional side chains or cross-linking. Table 7.2 Illustrates the inverse relationship between T_g and molecular weight. Additionally, Fig. 7.23 shows the change in T_g as a function of solution viscosity (as an indirect proxy for molecular weight) for a series of ethyl cellulose polymers.

T_g is an important fundamental property of amorphous and semicrystalline polymer systems, and it is frequently correlated with properties such as plasticity, ductility, film-forming ability, and stickiness. While polymer molecular weight is a major factor determining T_g values, numerous other structural factors will impact molecular rigidity and order, and, by implication, T_g. Some of these factors are summarized in Table 7.3.

Similarly to T_g, melting point temperatures and melt viscosity increase with increasing molecular weight. Operationally, polymer flow is typically characterized using a melt flow index, which is an inverse measure of melt viscosity. These properties are especially critical in modified-release dosage form design in the context of thermal processing, such as melt extrusion and injection molding. Frequently, higher melt flow index polymers are required for thermal processing in the range at or below 100°C in order to avoid thermal degradation of drug and acceptable throughput rates. Fig. 7.24 shows the variation of melt flow indices for different molecular weight grades of HPC, a thermoplastic polymer that is commonly used in melt extrusion processes, such as the manufacture of oral film strips and other extruded drug delivery devices.

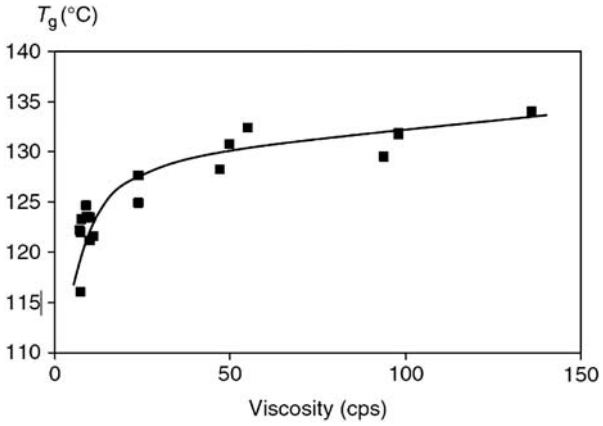


FIGURE 7.23 Glass transition temperature (T_g) as a function of solution viscosity (an indicator of molecular weight) for a series of Ashland ethyl cellulose polymers. In all cases, substitution was between 49.8 and 51 wt.% ethoxyl content.

TABLE 7.2 Different Molecular Weights of PVP and Their Glass Transition Temperatures (T_gs)

Product	K value	M_w[a]	T_g (°C)
Plasdone K12	10.2–13.8	4000	120
Plasdone K17	15.5–17.5	10,000	140
Plasdone K25	2.14–26	34,000	160
Plasdone K29/32	29–32	58,000	164
Plasdone K90	85–95	1,300,000	174

[a]*Weight average molecular weight.*

TABLE 7.3 Factors Affecting Glass Transition Temperature (T_g) Values

Molecular structural feature	Impact on T_g
Increased molecular weight	Higher
Increased symmetry	Higher
Increased length of flexible side chain	Lower
Addition of chain-stiffening units	Higher
Addition of polar groups	Higher
Increased cross-link density	Higher

7.2.6.1.6 Effect of molecular weight on gel strength

The relationship between gel strength and polymer molecular weight is of particular importance in the field of modified-release dosage form design. For hydrophilic matrix tablets, the gel strength will determine the erodibility of the matrix system. Erosion can

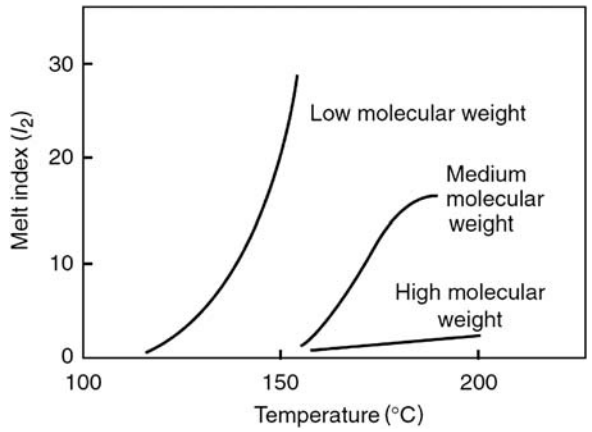


FIGURE 7.24 Melt index, I_2, as a function of temperature and Klucel™ hydroxypropyl cellulose molecular weight. The melt index is defined as the grams of molten material that is extruded under standard American Society for Testing and Materials (ASTM)-defined conditions in 10 min at a given temperature and applied load.

be described as polymer dissolution or the disentanglement of polymer chains from the gel surface and transfer of the polymer to the bulk solution. Erosion can be used to the formulator's advantage when designing a delivery system for insoluble drugs. Here, erosion is the main mechanism facilitating transfer of the insoluble drug out of the tablet matrix and into the dissolution medium. However, poor release characteristics such as variable burst release and dose dumping may be expected for a highly soluble drug that is formulated in a highly erodible dosage form.

For hydrophilic polymers such as hypromellose and HPC compressed into matrix tablets, the disentanglement concentration of the polymer follows a nonlinear, inverse relationship with molecular weight.[41] Furthermore, the polymer matrix erosion rate has also been shown to vary with the molecular weight in a nonlinear, inverse manner[42]:

$$\text{Erosion rate} = kM_n^{-a}$$

It should be noted that the opposite relationship applies to matrix swelling, that is, swellability increases with molecular weight up to a limiting threshold level. The rate and extent of polymer matrix swelling and erosion and the drug solubility and concentration are the dominant compositional factors determining drug-release kinetics from matrix tablet systems. Figs. 7.25–7.27 show the impact of the polymer molecular weight on the matrix tablet erosion, swelling, and drug release for a typical sparingly soluble compound, nifedipine.

7.2.6.2 *Side-chain substitution effects*

The structure of the side-chain substituents on the polymer backbone is a major compositional factor

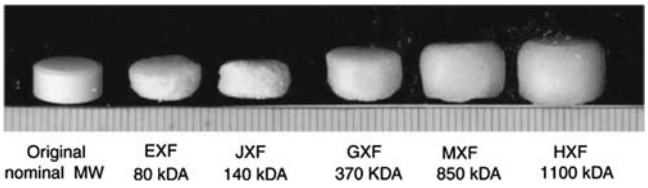


FIGURE 7.25 Effect of Klucel™ Pharm HPC molecular-weight grade on erosion and swelling behavior of nifedipine matrix tablets subjected to dissolution testing for 4 h. Formulation: 20% nifedipine, 30% HPC, 49.5% microcrystalline cellulose, and 0.5% magnesium stearate.

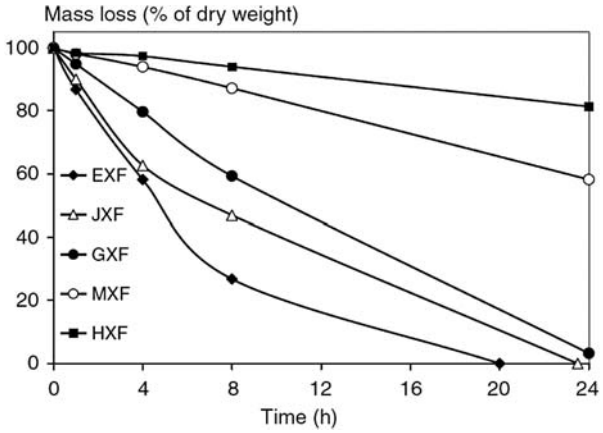


FIGURE 7.26 Effect of Klucel™ HPC molecular-weight grade on matrix erosion. Nominal molecular weights for Klucel™ EXF, JXF, GXF, MXF, and HXF Pharm HPC are 80, 140, 370, 850, and 1100 kDA, respectively. Formulation: 20% nifedipine, 30% HPC, 49.5% MCC, and 0.5% magnesium stearate.

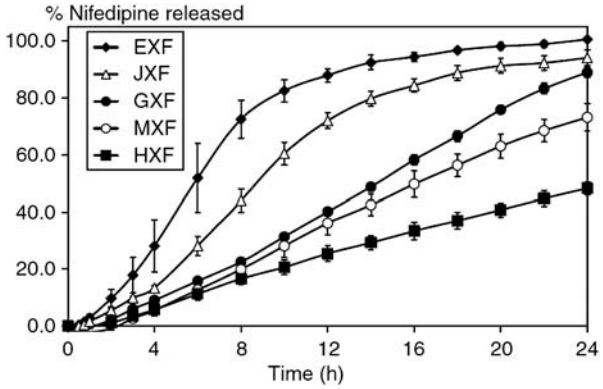


FIGURE 7.27 Effect of Klucel™ HPC molecular-weight grade on nifedipine dissolution. Nominal molecular weights for Klucel™ EXF, JXF, GXF, MXF, and HXF Pharm HPC are 80, 140, 370, 850, and 1100 kDA, respectively. Formulation: 20% nifedipine, 30% HPC, 49.5% MCC, and 0.5% magnesium stearate.

impacting polymer functionality. Important aspects of substitution are the chemical structure of the substituents, the extent of backbone substitution, and the uniformity of substitution. In this section, we will mainly focus on the impact of substituent type and the extent of substitution on the functionality in modified-release systems.

7.2.6.2.1 Side-chain structure (substituent type)

Effect of side-chain structure on polymer solubility. In most cases, substitution of less polar groups for hydrogens or hydroxyls on a polymer chain leads to a reduction in crystallinity and, usually, also a reduction in melting points. Such changes are thus generally expected to improve thermoplasticity and polymer solubility. A good example is presented by cellulose, a naturally occurring polymer comprised of anhydroglucose units. Native cellulose has considerable microcrystalline character. When the free hydroxyls on the anhydroglucose backbone are substituted with hydroxyalkyl side chains, for example, hydroxypropyl or hydroxyethyl groups, the crystallinity of the resultant hydroxypropyl or hydroxyethyl cellulose is below the limit of reliable quantification (<10%). Moreover, unlike microcrystalline cellulose, the derivatized cellulose ethers are freely gelling and soluble in water. Polymer water solubility can be further enhanced by deliberate inclusion of highly polar, ionizable substituents. Sodium carboxymethyl cellulose (NaCMC) is an example of this.

Unlike small drug molecules, it is not useful to describe the solubility of a polymer in a given solvent system in terms of saturation concentration values (eg, g/mL). Among the more useful methods, solubility parameters can be used to evaluate the solubility and compatibility between a polymer and solvent or other additives with solvent-like properties. Additionally, for water solubility, equilibrium moisture content (hygroscopicity) and cloud point are useful indirect measures of water solubility. The solubility parameters for a series of differently substituted cellulose ether molecules and other materials of interest to formulators are shown in Table 7.4.

For many polymers including cellulose, the relationship between water solubility and the nature of the side chain or substituent chemistry can be understood in terms of hydrogen bonding between the hydrogen atoms of water and the polar oxygen atoms present in the polymer backbone and on the side chains. Studying a series of alkali celluloses of similar molecular weight, Klug[43] showed that polymer hydrophilicity increases as the oxygen content of the derivatized cellulose increases due to either the type or the extent of substitution. Sodium carboxymethyl cellulose is an exception due to its ionic character. The increased hydrophilicity impacts numerous important properties such as hygroscopicity, cloud point, matrix tablet swelling ability, and matrix gel strength. The general relationship is shown in Fig. 7.28.

TABLE 7.4 Solubility Parameters for a Series of Cellulose Ethers and Other Materials of Interest

Polymer or drug molecule	Average solubility parameter,[a] δ ($MPa^{½}$)
EC (2.7 DS)	19.2
HPC (3.8 MS)	23.2
HPMC (1.9 MS)	23.6
HEC (2.5 MS)	25.0
NaCMC (0.7 DS)	28.9
Cellulose	31.1
PVA	30.2
PEO	19.3
Metformin HCl (highly water-soluble drug)	27.2
Itraconazole (low water-soluble drug)	19.5

[a]*Calculated as the Average of the Hildebrand, Hansen, Fedors, and van Krevelen Solubility Parameters.*

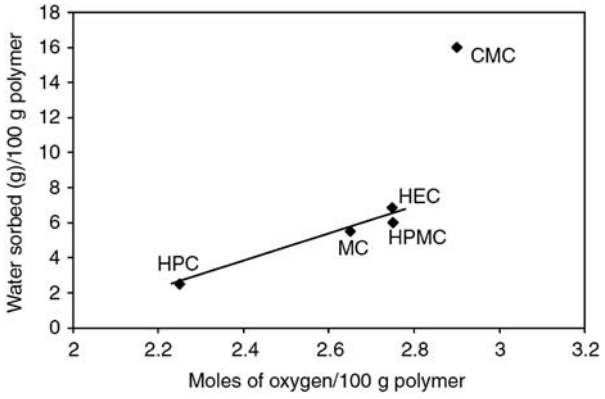


FIGURE 7.28 Effect of oxygen content on the equilibrium moisture content of a series of cellulose derivatives. *HPC*, hydroxypropyl cellulose; *MC*, methyl cellulose; *HPMC*, hydroxypropyl methylcellulose; *HEC*, hydroxyethyl cellulose; *CMC*, sodium carboxymethyl cellulose. *Adapted from Klug, E.D. Some properties of water soluble hydroxyalkyl celluloses and their derivatives. J Polym Sci Part C 1971;36:491–508.*

Effect of side-chain structure on gel strength, swelling, and erosion in matrix systems. From Fig. 7.28, it can be seen that the differently substituted cellulose derivatives can be ranked in the following order of ascending hydrophilicity: HPC < MC < HPMC < HEC and CMC. The difference in hydrophilicity due to substitution on the polymer backbone is particularly important for the functionality of hydrophilic matrix systems. As shown in Fig. 7.29, polymer hydrophilicity, in combination

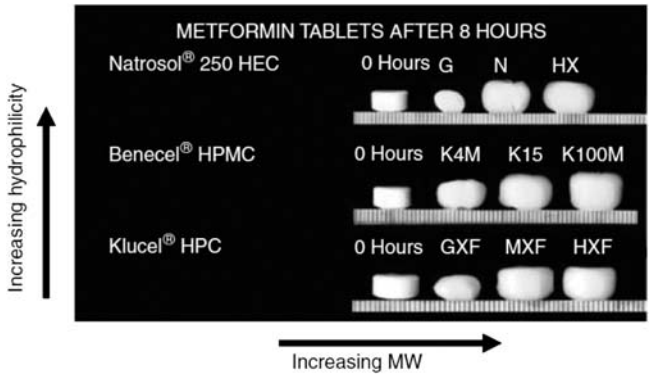


FIGURE 7.29 Matrix tablet swelling and erosion behavior as a function of hydrophilicity (cellulose ether substitution) and molecular weight.

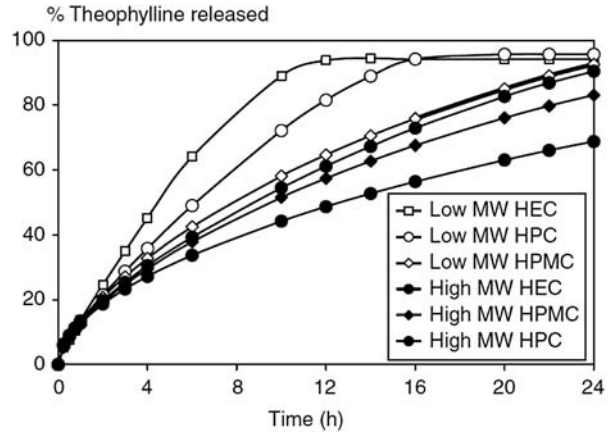


FIGURE 7.30 Drug release from matrix tablets as a function of cellulose ether chemistry and molecular weight.

with molecular weight, has a major impact on matrix swelling and erosion behavior. For medium and low-soluble drugs, these factors are the main determinants of release mechanism and release kinetics (Fig. 7.30).

Effect of side-chain structure on mechanical properties. The nature of the side-chain substituent type also significantly impacts mechanical and thermoplastic properties. HPC and ethyl cellulose (EC) are significantly more thermoplastic than HPMC and HEC. The compactibility of matrix-forming cellulose ethers has been reported to increase in the following rank order: HEC < HPMC < HPC.[44]

7.2.6.2.2 Extent of side-chain substitution

In addition to the identity of the side-chain structure, the amount of side-chain substitution on the polymer backbone can also exert a significant effect on physical and chemical properties of the polymer.

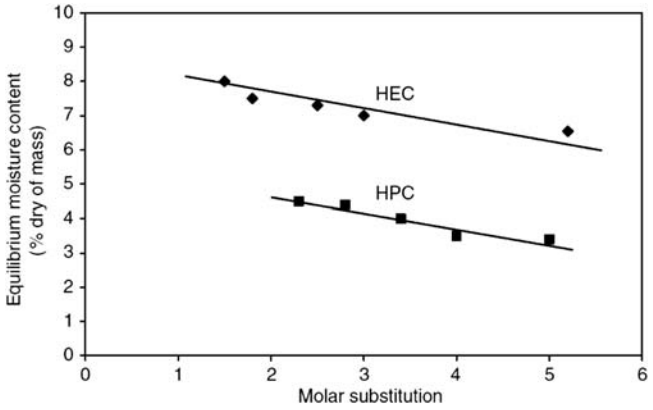


FIGURE 7.31 Effect of substitution level on the equilibrium moisture content of hydroxypropyl and hydroxyethyl cellulose at 50% relative humidity. *Adapted from Klug, E.D. Some properties of water soluble hydroxyalkyl celluloses and their derivatives. J Polym Sci Part C 1971;36:491–508.*

7.2.6.2.3 Effect of extent of substitution on solubility

When highly polar hydroxyl groups on crystalline cellulose are substituted with hydroxyalkyl groups to manufacture HPC or HEC, water solubility initially increases due to a reduction in crystallinity and hydrogen bonding between the cellulose backbone chains. However, as the amount of hydroxyalkyl substitution continues to increase, the polymer becomes increasingly hydrophobic. As shown in Fig. 7.31, the equilibrium moisture content steadily decreases as MS increases from 2.0 to 5.0 for both HEC and HPC. A similar relationship has also been demonstrated for the cloud point.[43] An exception to this behavior is polymers with ionic groups in their side chains. In this case, increasing the level of highly polar substituents will increase water solubility. For example, when the DS for sodium carboxymethyl cellulose is increased from 0.7 to 1.2, the equilibrium moisture content at 50% relative humidity increases from 13% to 18%.[43]

7.2.6.2.4 Effect of extent of substitution on amorphous solid dispersion properties

As explained, the amount of side-chain substitution or the ratio of different substituents will affect the relative hydrophilic-hydrophobic balance of the polymer. When using polymers as carriers for amorphous solid dispersion of poorly soluble drugs, this has the potential to affect the polymers' ability to undergo hydrogen bonding with drug molecules in the solid state, as well as to undergo hydrophobic interaction with drug molecules in the saturated solution state. For example, as shown in Fig. 7.32, HPMCAS is available in three different acetyl/succinyl substitution ratios.

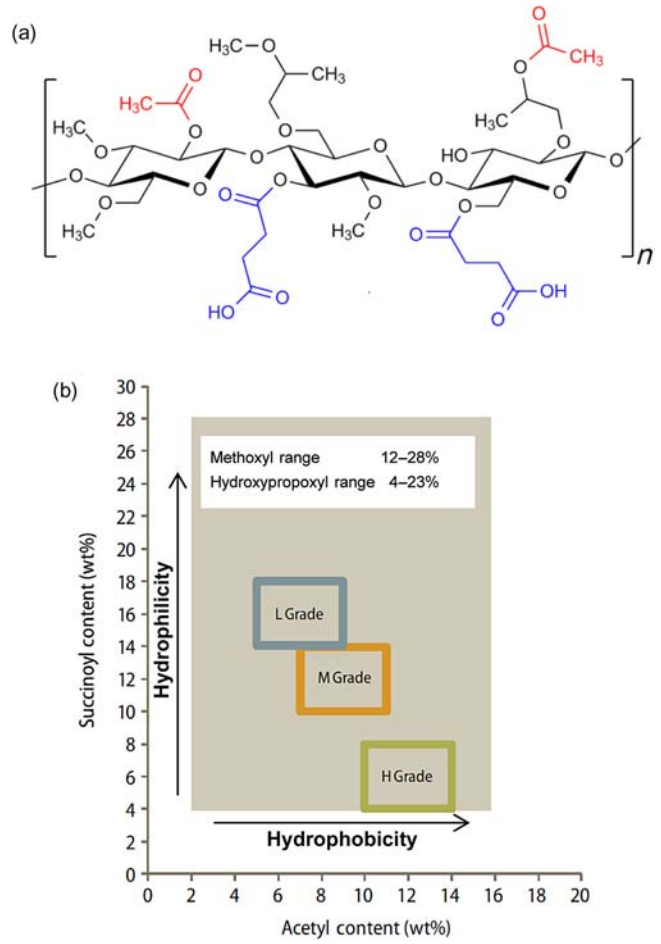


FIGURE 7.32 (a) Idealized chemical structure of HPMCAS. (b) Schematic showing acetyl and succinoyl substitution ranges and effect on hydrophilic/lipophilic balance for different commercial grades of Aquasolv™ HPMCAS.

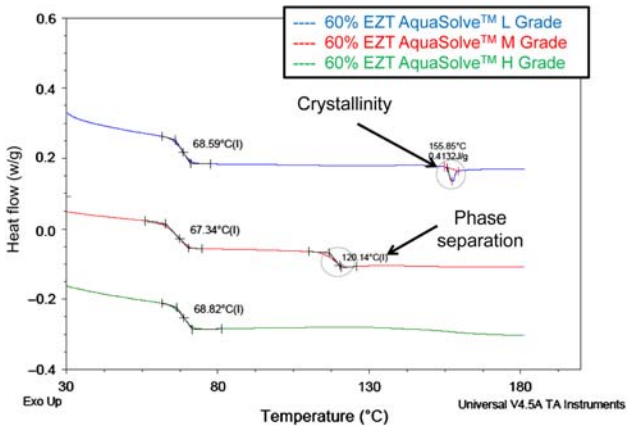


FIGURE 7.33 DSC thermograms for ezetimibe: Aquasolv™ HPMCAS solid dispersions aged for 65 h in open dish conditions at 40°C and 75% relative humidity.

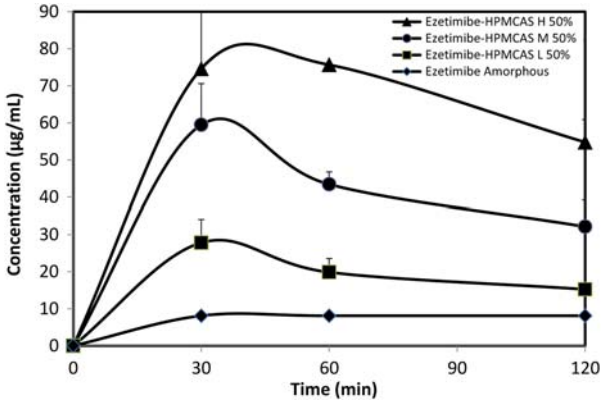


FIGURE 7.34 Nonsink condition dissolution profiles of amorphous ezetimibe and ezetimibe-Aquasolv™ HPMCAS solid dispersions equivalent to 500 µg/mL ezetimibe in 20 mL FaSSIF media at 37°C and 300 rpm using fiber-optic dissolution profiler, µDISS Profiler. At 30, 60 and 120 min, samples ($n = 4$) were centrifuged at 3700 rpm for 5 min, and the supernatant was analyzed for drug concentration.

These different substitution grades can stabilize amorphous solid dispersions to varying degrees, depending on drug properties. Fig. 7.33 shows that H grade is significantly better at preventing recrystallization of rapid recrystallizer ezetimibe in the solid amorphous state than M and L grade, both of which have higher succinoyl content. This can be related to the ability of HPMCAS H grade to undergo better hydrogen bonding.[45]

In addition, the acetyl/succinyl substitution level also affects precipitation inhibition in saturated drug solutions that occurs after the amorphous dispersion is dissolved. Fig. 7.34 shows that HPMCAS with high acetyl, that is, greater hydrophobicity, is more effective at preventing ezetimibe precipitation from saturated solutions. This can be related to an improved hydrophobic interaction and drug-polymer binding capability.[46]

7.2.6.2.5 Effect of extent of substitution on mechanical properties

In addition to increased hydrophobicity, an increase in the amount of less polar substituents generally also results in an increase in polymer plasticity. As shown in Fig. 7.35, in the case of tensile film strength, this manifests in lower ultimate tensile strength but greater elongation and film flexibility. In contrast, increasing the amount of highly polar, ionic side chains tends to result in an increased T_g, a markedly less thermoplastic material. As a result, films will tend to exhibit relatively high tensile strength but will be very brittle and inflexible, with minimal elongation at break. Sodium carboxymethyl cellulose shows this behavior and, as a result, requires high levels of added plasticizer if useful, defect-free films suitable for tablet or particle coating are to be achieved.

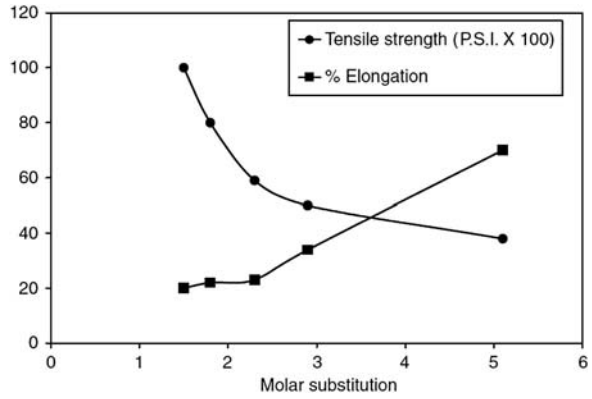


FIGURE 7.35 Effect of amount of substitution on the tensile strength and percentage elongation of a hydroxyethyl cellulose film at constant thickness, 50% relative humidity and 25°C. *Adapted from Klug, E.D. Some properties of water soluble hydroxyalkyl celluloses and their derivatives. J Polym Sci Part C 1971;36:491–508.*

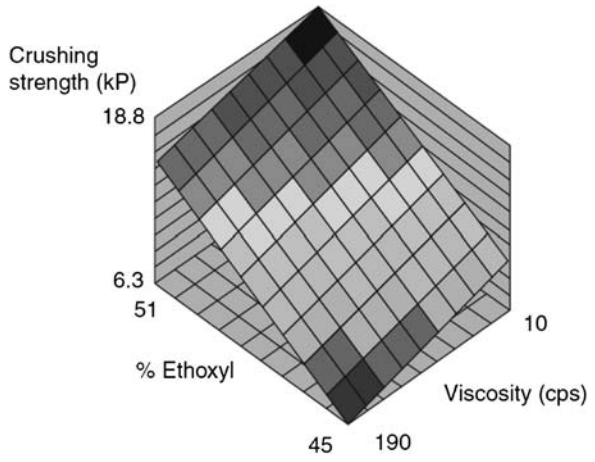


FIGURE 7.36 Effect of ethoxyl content and viscosity (molecular weight) on compactibility of ethyl cellulose (EC). Maximum tablet strength is achieved for high ethoxyl, low viscosity EC.

TABLE 7.5 Tablet Tensile Strength for a Series of Hypromellose and Methyl Cellulose Samples of Similar Viscosity When Compressed at 15 kN[47]

Polymer type	Methoxy (%)	Hydroxy propyl (%)	Tensile strength (MPa)
HPMC 2208	10–24	4–12	1.5
HPMC 2910	28–30	7–12	1.6
HPMC 2906	27–30	4–7.5	2.8
MC	27.5–31.5	0	3.5

TABLE 7.6 Effect of Ethoxyl Content and Viscosity on Selected Solid-State Properties of Ethyl Cellulose

Ethoxyl (wt.%)	Viscosity (cps)	Crystallinity (%)	Melting point (°C)	T_g (°C)	Crushing force (kP)[a]
50.4	9	24.6	257.0	122.1	20.9
50.8	9	28.5	261.1	124.6	18.8
50.0	50	15.3	246.6	130.7	14.6
49.6	94	17.8	261.0	129.5	16.1
48.0	10	9.1	210.0	131.0	14.7
48.5	94	7.9	224.3	133.5	11.5
47.5	10	8.2	178.6	135.1	12.3

[a]*275 mg pure ethyl cellulose tablets compressed on a Manesty BetaPress at 25 kN.*

For polymers compressed into tablets, increasing the number of nonpolar substituents generally results in increased plasticity, lower elastic recovery after compression and ejection, and consequently denser compacts with higher diametral crushing strength (tensile failure). Similarly, decreasing the number of more polar substituents will have the same effect. The compaction behavior of a series of hypromellose and methyl cellulose grades with varying levels of methoxyl (less polar, hydrophobic substituent) and hydroxypropyl (more polar, more hydrophilic substituent) groups illustrates this.[47] As can be seen in Table 7.5, as the percentage of methoxyl in the polymer increases and the percentage of hydroxypropyl decreases, for a series of similar viscosity polymers, the tablet strength significantly increases.

An additional example of the significant effect that an increase in the number of nonpolar substituents can have is provided by ethyl cellulose (EC). It has been reported that a higher compactability of EC was associated with high ethoxyl content and low molecular weight (low viscosity) (Fig. 7.36).[92]

High substitution and low molecular weight were also correlated with significantly higher crystallinity and higher melting points, while the amorphous regions had lower glass transition temperatures (Table 7.6). Compaction simulator studies showed that the net effect of this unique solid-state structure is a marked reduction in postcompaction elastic recovery of the polymer (Fig. 7.37). Less energy of compaction was therefore lost due to postcompaction axial expansion, resulting in denser compacts with lower porosity. This had a significant effect on drug diffusion from nonswelling, porosity-controlled matrix tablets. Additionally, the reduced viscoelasticity resulted in lower strain rate sensitivity (10% vs typically 20–25% for other EC types).

7.2.6.3 Copolymerization

7.2.6.3.1 Thermal properties of copolymers

When the monomers of two crystalline homopolymers are combined, the degree of crystallinity and the

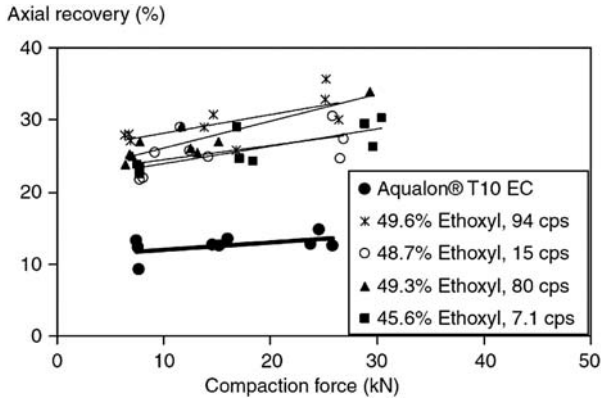


FIGURE 7.37 Effect of ethoxyl content and viscosity on postcompaction axial recovery. High ethoxyl, low-viscosity Ashland T10 EC has the lowest axial recovery.

melting point are usually depressed. The following relationship frequently applies:

$$\frac{1}{T_m} = \frac{1}{T_m^0} - \frac{R}{\Delta H_m}\ln(n)$$

where:

n is the mole fraction of crystalline constituent;
T_m is the melting point of the copolymer;
T_m^0 is the melting point of the pure homopolymer;
ΔH_m is the heat of fusion.[2]

In contrast, the glass transition temperature, T_g, of random copolymers usually falls into the range between that of the two corresponding homopolymers. This difference between how T_m and T_g of a copolymer are affected can be attributed to the fact that crystal structure disruption is not relevant to T_g of the amorphous polymer domains. Frequently, a simple weight-average mixing rule of the following type can be applied to get an approximate idea of the resultant T_g of the copolymer:

$$a_1 w_1 (T_g - T_{g1}) + a_2 w_2 (T_g - T_{g2}) = 0$$

where:

T_{g1} and T_{g2} refers to the individual homopolymers;
w_1 and w_2 are weight fractions of monomers 1 and 2;
a_1 and a_2 depend on monomer type.[48]

The 60:40 random copolymer of vinylpyrrolidone and vinyl acetate (copovidone) is a good example of this phenomenon. As shown in Fig. 7.38, the addition of vinyl acetate to vinylpyrrolidone results in a T_g of 105°C, which intermediate between polyvinylpyrrolidone (povidone) of similar molecular weight (165°C) and polyvinyl acetate (70°C) and a significant T_g reduction when compared to polyvinylpyrrolidone alone.

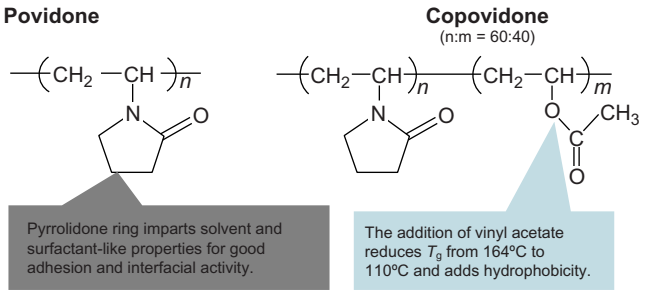


FIGURE 7.38 Schematic structures of povidone and copovidone.

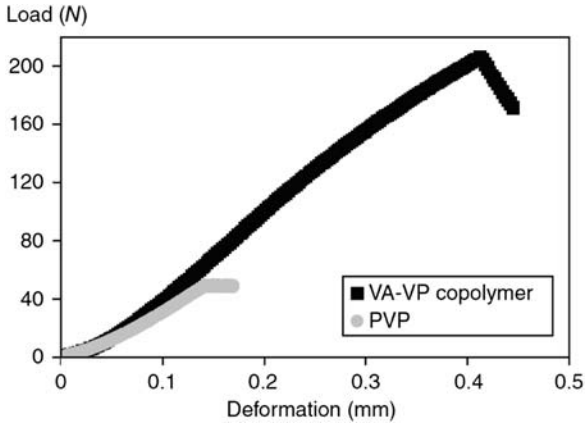


FIGURE 7.39 Load–deformation plots for the pure polymer tablets of povidone ($T_g \sim 161$°C and copovidone ($T_g \sim 101$°C) subjected to diametral compression.

7.2.6.3.2 Mechanical properties of copolymers

The decrease in T_g, and T_m as a result of adding a comonomer to an amorphous or crystalline generally results in an increase in copolymer plasticity and flexibility. Therefore, copolymers are frequently better film formers and tablet binders than the initial homopolymer. Povidone (PVP) and copovidone, the 6:4 random copolymer of vinylpyrrolidone and vinyl acetate (VP–VA copolymer) present a relevant pharmaceutical example. PVP, although frequently used as a tablet binder, is relatively stiff, brittle, and hygroscopic. The brittleness and stiffness are reflected in the relatively high T_g of approximately 164°C. In contrast, the VP–VA copolymer has a significantly lower T_g of approximately 110°C, is a lot more flexible and plastic, and has lower moisture uptake than PVP. This results in a significantly improved compactibility, as seen in Fig. 7.39.

7.3 COMMONLY USED POLYMER EXCIPIENTS IN SOLID ORAL PRODUCTS

Polymers are widely used as excipients or sometimes even active ingredients in pharmaceutical applications. The selection of excipients for a particular

application will depend on the therapeutic target, route of administration, dosage forms, and drug-release profile, in addition to the cost, availability, and complexity of related pharmaceutical unit operations.

The basic architecture of polymeric excipient materials to be considered consists of a relatively simple linear polymer structure. This includes cellulosic derivatives and most major synthetic polymers used as formulation components. Aside from a limited number of cross-linked polymers such as croscarmellose, lightly cross-linked poly (acrylic acid) (Carbopol), and crospovidone, most polymeric materials used as excipients are not branched or cross-linked. However, a number of materials do participate in reversible self-association that can result in the formation of dynamically branched or physically cross-linked system, and this secondary association equilibrium can exert an influence on the usage properties of the material.

Polymeric excipients are numerous and very diverse of natural, semisynthetic, or synthetic origin, depending on the source and method of preparation. They can be classified according to their source of origins: natural polymers and synthetic polymers. They can also be classified alternatively according to their chemical composition, chain shape, comonomer sequence, and their performance.

Biopolymers or natural polymers refer to polymers of natural origin, such as polysaccharides. They are produced by living organisms or from biomass derived from monomers such as sugars for polysaccharides. Natural polymeric materials such as shellac, amber, and natural rubber have been in use for centuries. A variety of other natural polymers exist, such as cellulose and hydrocolloids. Natural polymers can also be chemically modified to produce semisynthetic polymers. Commercially important semisynthetic polymers include ethyl cellulose (EC), hydroxypropyl cellulose (HPC), hydroxypropyl methylcellulose (HPMC), hydroxypropyl methylcellulose phthalate (HPMCP), hydroxypropyl methylcellulose acetate succinate (HPMCAS), and methyl cellulose (MC) for pharmaceutical applications.

Synthetic polymers are man-made by chemical synthesis from monomers. Polyacrylates, polyvinylpyrrolidone (PVP), polyvinyl alcohol (PVA), poly(lactic-*co*-glycolic acid) (PLGA), poly(ethylene glycol) (PEG), and polyethylene (PE) are typical examples of synthetic pharmaceutical polymers. Synthetic and semisynthetic polymers have a potential advantage over natural polymers due to their consistency and reproducibility in quality and functional properties.

Rather than presenting a comprehensive coverage of the molecular properties of all classes of polymeric excipients used in pharmaceutical formulations, we will focus our attention primarily on an overview of those typically used in solid oral drug products. These polymeric excipients are used to fulfill a variety of functions in solid oral dosage forms such as binders, disintegrants, coatings, and drug-release, rate-controlling polymers.

The chemical structure and applications of a number of important pharmaceutical polymer excipients for oral dosage forms are shown in Table 7.7.

7.3.1 Cellulose and cellulose derivatives

Cellulose is abundant in nature and exists in plant-based foods. It is a safe material for oral drug formulations based on a long history of human consumption. Cellulose and its derivatives (ether and ester) (Fig. 7.40) are among the excipients most frequently used in solid oral dosage forms. A number of overviews of polysaccharide chemistry in general,[49] and cellulose derivatives in particular,[50] are readily available, and these classic references remain worthwhile resources towards understanding key aspects of the chemistry.

The simplest modification of cellulose is size reduction. Powdered cellulose is produced by the purification and mechanical size reduction of α-cellulose obtained as a pulp from fibrous plant materials. It occurs as a white or almost white, odorless, and tasteless powder of various particle sizes, ranging from a free-flowing fine powder or granular dense powder to a coarse, fluffy, nonflowing powder. Powdered cellulose is used as a tablet diluent and filler in two-piece hard capsules.[51,52] Powdered cellulose has acceptable compression properties. Low-crystallinity powdered cellulose has exhibited properties that are different from standard powdered cellulose materials, and it has shown potential as a direct-compression excipient. In soft gelatin capsules, powdered cellulose may be used to reduce the sedimentation rate of oily suspension fills. It is also used as the powder base material of powder dosage forms and as a suspending agent in aqueous suspensions for peroral delivery. It may also be used to reduce sedimentation during the manufacture of suppositories.

In order to improve the compactability of cellulose and make it useful as dry binder for roller compaction and tableting by direct compression, cellulose is modified by controlled hydrolysis with dilute mineral acid solutions of α-cellulose, obtained as a pulp from fibrous plant materials. Following hydrolysis, the hydrocellulose is purified by filtration, and the aqueous slurry is spray dried to form dry, porous particles of a controlled size distribution and moisture content for different applications. This process converts cellulose to microcrystalline cellulose (MCC). MCC is purified, partially depolymerized cellulose that occurs

TABLE 7.7 List of Polymeric Excipients for Oral Dosage Forms

Polymers	Major functionality	Applications	Typical use levels (%)	IID or IIG limit[a]
WATER-INSOLUBLE NATURAL/MODIFIED POLYMERS				
Cellulose, microcrystalline (MCC) $(C_6H_{10}O_5)_n$, $n \sim 220$	Binder/filler	Tablet and capsules	10–90	789.6 mg
Cellulose, powdered $(C_6H_{10}O_5)_n$	Filler	Tablet and capsules	5–40	391.7 mg
Cellulose, silicified, microcrystalline (SMCC), $(C_6H_{10}O_5)_n$. SiO_2	Binder/filler	Tablet and capsules	10–90	N/A
Cellulose acetate (CA) $(C_6H_7O_2(OH)_3)_n$	Coating for osmotic pump	Osmotic pump	1–20	47.49 mg
Ethyl cellulose (EC) $C_{12}H_{23}O_6(C_{12}H_{22}O_5)_nC_{12}H_{23}O_5$	Coating/binder	Extended-release coating/ matrix former	1–30	308.8 mg
Croscarmellose sodium (NaCMC) $(C_8H_{16}O_8)_n$	Disintegrant	Tablet and capsules	0.5–25	180 mg
Sodium starch glycolate (SGS) $(C_{24}H_{44}O_6N_a)_n$	Disintegrant	Tablet and capsules	2–8	876 mg
WATER-INSOLUBLE SYNTHETIC POLYMERS				
Crospovidone $(C_6H_9NO)_n$	Disintegrant	Tablet and capsules	2–30	340 mg
Polyacrylic acid (carbomer; Carbopol), $(C_3H_4O_2)_n$	Rheology modifier/controlled release	Tablets, suspension	0.1–30	90 mg
Polymethacrylate, cationic	Coating/binder	Extended-release coating/ matrix former	5–50	161 mg
Polymethacrylate, neutral	Coating/wet binder	Extended-release coating/ matrix former	5–50	187.3 mg
WATER-SOLUBLE NATURAL/MODIFIED NATURAL POLYMERS				
Hydroxyethyl cellulose (HEC) $[C_6H_7O_2(OH)_x(OC_2H_5)_y[O(CH_2CH_2O)_mH]_z]_n$	Coating/rheology modifier	Tablet, suspension, pellets	2–50	400 mg
Hydroxypropyl cellulose (HPC) $(C_{36}H_{70}O_{19})_n$	Binder/coating/controlled-release matrix, extrusion aid	Tablet, capsules, pellets	1–50	240 mg
Hydroxypropyl methylcellulose (HPMC or hypromellose) $(C_{56}H_{108}O_{30})_n$	Binder/coating/controlled-release matrix	Tablet, capsules, pellets	1–75	480 mg
Methyl cellulose (MC) $C_6H_7O_2(OH)_x(OCH_3)_y$	Coating/binder/rheology modifier	Tablet, suspension, capsules	0.5–20	30 mg
Sodium alginate $(C_6H_7O_6Na)_n$	Coating/controlled-release matrix/ rheology modifier	Tablet, capsules, pellets	1–40	350 mg
Sodium carboxymethyl cellulose (NaCMC) $(C_8H_{15}NaO_8)_n$	Binder/coating/controlled-release matrix/rheology modifier	Tablet, capsules, pellets	0.1–90	2000 mg
WATER-SOLUBLE SYNTHETIC POLYMERS				
Polyethylene glycol (PEG), $C_{2n}H_{4n+2}O_{n+1}$	Plasticizer, solubility enhancer	Coating, tablet, softgel, capsules	1–40	960 mg
Polyethylene oxide (PEO), $C_{2n}H_{4n+2}O_{n+1}$	Mucoadhesion, tablet binder, matrix former, thickener	Tablets, coating	5–50	393.46 mg
Polyvinyl alcohol (PVA), $(C_2H_4O)_n$	Film former, thickener	Tablets, coating, microspheres	0.5–20	79.4 mg
Polyvinylpyrrolidone (PVP), $(C_6H_9NO)_n$	Coating, binder, solubility enhancement	Tablet, capsules, pellets	0.5–90	853.8 mg
POLYMERS WITH pH-DEPENDENT WATER SOLUBILITY				
Cellulose acetate phthalate (CAP)	Enteric coating	Tablet and capsules	0.5–9	75.6 mg
Hydroxypropyl methylcellulose acetate succinate (HPMCAS)	Enteric coating, solubility enhancement	Tablet and capsules	0.5–90	560 mg
Polymethacrylate, anionic	Enteric coating, solubility	Enteric coating, solubility	0.5–30	430.8 mg

[a]*For the latest IID or IIG Limit: http://www.accessdata.fda.gov/scripts/cder/iig/index.Cfm.*

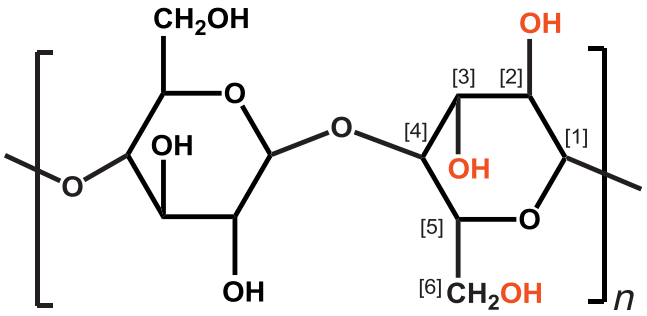


FIGURE 7.40 Basic structure of cellulose, which is used as the starting material for any cellulosic derivative. Each anhydroglucose residue possesses three potentially reactive hydroxyl groups. These are indicated and numbered in the second residue according to the associated carbon on which they reside. For most reaction schemes, the 2 and 6 positions are more reactive than the 3 position.

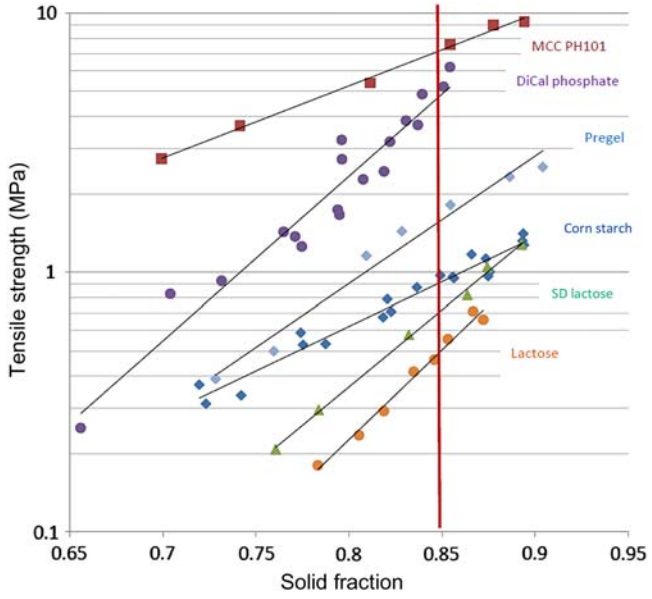


FIGURE 7.41 Comparison of tensile strength of tablets made from different materials at a solid fraction of 0.85, an average value for commercial tablets. *Adapted from Amidon, G.E. Compaction and the properties of mixtures, presented at Compaction Simulation Forum, Cambridge, MA, November 13–14, 2012, http://people.clarkson.edu/~ccetinka/CompactionPresentation/1%20Amidon%20UMichigan.pdf.*[61]

as a white, odorless, tasteless, crystalline powder composed of porous particles. The properties and applications of MCC depend on the particle sizes and moisture grades of MCC.[53–55] MCC is widely used in pharmaceuticals, primarily as a binder/diluent in oral tablet and capsule formulations where it is used in both wet-granulation and direct-compression processes. In addition to its use as a binder/diluent, MCC also has some lubricant and disintegrant properties that make it useful in tableting. A comparison of tensile strength of tablets made from different materials at various solid fractions is shown in Fig. 7.41 where it can be seen that microcrystalline cellulose is the most compactable filler. Therefore, it is not surprising that MCC has been a widely used excipient for tablets.[56–60]

Cellulose can also be mixed with other ingredients to form coprocessed excipients. One of the most widely used coprocessed cellulose is silicified microcrystalline cellulose (SMCC).[62–64] Silicified microcrystalline cellulose is a synergistic, intimate physical mixture of two components: microcrystalline cellulose and colloidal silicon dioxide. It is manufactured by cospray drying a suspension of microcrystalline cellulose particles and colloidal silicon dioxide such that the dried finished product contains 2%w/w colloidal silicon dioxide. The colloidal silicon dioxide appears physically bound onto the surface and inside the silicified microcrystalline cellulose particles. Extensive studies using different spectroscopic methods have not detected any form of chemical interaction.

Cellulose has also been coprocessed with lactose, mannitol, and sodium carboxymethyl cellulose for applications in direct compression for tablets as binder/filler and in oral suspensions as a suspension vehicle.

In order to improve the chemical and physical properties of cellulose for food and pharmaceutical applications, chemical modifications have been applied to cellulose to produce a diverse range of cellulose derivatives for pharmaceutical applications.

As the name implies, cellulose derivatives are all based on cellulose provided by a variety of starting furnishes (wood pulp, chemical cotton). The starting biopolymer, cellulose, is a β-1–4 linked linear polymer of anhydroglucose (Fig. 7.40). Due to the relatively stiff main chain backbone, the array of available hydrogen bonding sites per monomer residue, and structural order inherent in the natural system, cellulose molecules form crystalline microfibrils that are mechanically strong and highly resistant to enzymatic attack. Cellulose is, therefore, insoluble in water and indigestible in the human body. Dissolution of cellulose normally requires exotic solvent systems or in situ derivatization to cap self-associating hydroxyl groups.

The reaction chemistry used in the production of cellulosic derivatives[50] is based on removing that hydrogen bond-mediated order from the system via activation of purified cellulose with a strong caustic, followed by reaction with appropriate alkylating agents (typically alkyl halides) and/or etherifying agents (eg, propylene oxide or ethylene oxide) to introduce various hydrophilic and/or hydrophobic substituents. This derivatization reaction introduces point-packing defects along the main backbone and prevents the reestablishment of the higher level of order that was present in the native cellulose. The overall result is that the material is rendered soluble in a

TABLE 7.8 Structure of Cellulose and Its Derivatives Based on Fig. 7.40

Cellulose derivatives	*R* groups
Cellulose	–H
Methyl cellulose (MC)	–H and $-CH_3$
Ethyl cellulose (EC)	–H and $-CH_2CH_3$
Hydroxyethyl cellulose (HEC)	–H and $-CH_2CH_2OH$
Hydroxypropyl cellulose (HPC)	–H and $-CH_2CHOHCH_3$
Hydroxypropyl methylcellulose (HPMC)	–H and $-CH_3$ and $CH_2CHOHCH_3$
Hydroxypropyl methylcellulose acetate succinate (HPMC AS)	–H and $-CH_3$ and $CH_2CHOHCH_3$, $-C(O)CH_3$, $-C(O)CH_2CH_2COOH$
Cellulose acetate	–H, $-C(O)CH_3$
Cellulose acetate phthalate	–H, $-C(O)CH_3$, $-C(O)C_6H_4COOH$
Sodium carboxymethyl cellulose (Na CMC)	–H and $-CH_2COONa$

variety of solvents, with the preferred solvents being dependent on the substituents bound to the backbone. The overall molecular weight of the cellulose derivative is largely controlled by the specific selection of starting cellulose furnish for the higher-molecular-weight grades, with either in situ or post-derivatization oxidative or hydrolytic glycosidic chain scission being used to prepare the lower-molecular-weight members of a given chemical class (Table 7.8).

Much of the solution chemistry of cellulose derivatives can be organized by understanding the interplay of the substituent as a defect in the native cellulose packing structure to increase solubility with the use of substituent hydrophobicity and/or charge to tailor compatibility with nonaqueous or binary water/nonaqueous solvents.

Cellulose derivatives can be further classed according to whether or not the substituents can undergo a chain extension oligomerization (either in free solution or once grafted onto the cellulose backbone). Hydroxypropyl and hydroxyethyl substituents are able to undergo chain extension through the inherent hydroxyl group, whereas carboxymethyl, methyl, and ethyl substituents effectively cap the reactive hydroxyl site when they are grafted onto the chain. In cases where the substituent is able to undergo chain extension, the degree of substitution (DS = number of substituent capped hydroxyls per anhydroglucose) and total molar substitution (MS = number of substituent moles per mole of anhydroglucose) will differ with MS > DS. Note, for any cellulose derivative, the maximum value for DS cannot exceed three, as each anhydroglucose unit only has three hydroxyl groups available for reaction (ie, DS < 3), whereas the MS can be greater than 3. For example, for most water-soluble cellulose ethers, the DS values are in the range of 0.4–2, whereas for water-insoluble polyether such as ethyl cellulose, the DS values are between 2.3 and 2.8. On the other hand, typical MS values for hydroxyalkyl ethers of cellulose are between 1.5 and 4.0.

Cellulose derivatives are generally graded according to their viscosity under defined concentration conditions, which is ultimately tied to their molecular weight and substitution levels as expressed by the DS, MS, or percent mass of grafted substituent. Note that in some cases (ethyl cellulose would be a specific example), the substituent used to quantify the mass loading of derivatizing reagent on the polymer, $-OCH_2CH_3$, reflects the combination of the oxygen atom from anhydroglucose with the ethyl group from the ethyl chloride used in the derivatization reaction. Typical data of the various grades of cellulose ethers shown in Table 7.9 indicate that the range in average molecular weight represented in these polymers is extremely large—ranging from 50,000 Da on the low side to 1,200,000 Da for the highest molecular weight materials available. This is a factor of 24 in net molecular weight. Owing to the variable content in residue formula weight provided by the pendant groups, a direct comparison of the differences in the DP will differ slightly. In this case, net degrees of polymerization, basically the number of discrete monomer residues on average per chain, range from a low of roughly 215 for low-molecular-weight sodium carboxymethyl cellulose to a high of about 4700 for the highest molecular weight grade of hydroxyethyl cellulose. Again, the total dynamic range that chain length varies is by about a factor of 20.

The broad range in the average molecular weight, or equivalently, the DP, tells only part of the story. These ranges reflect the wide variability in the population average values. Within a given population, in other words, for a given molecular weight, there is also a very disparate population of overall chain lengths present. For example, in a size exclusion chromatographic characterization of any of the high-molecular-weight cellulosics, the net range of molecular weights encompassed by the population of a given moderate- to high-molecular-weight sample will typically span a range from a few hundreds to thousands up to a few million. In other words, the net diversity in chain lengths is extremely broad and easily spans a net range of two to three orders of magnitudes as the high- and low-molecular-weight portions of the distribution are accounted. A typical example is shown in Fig. 7.42. More typically, the polydispersity of a polymeric material is quantified by the ratio M_w/M_n. In cellulose

TABLE 7.9 Viscosity Grade Molecular Weights for Various Cellulose Derivatives

Viscosity grade (cP, at listed concentration)	Concentration (wt.%)	Viscosity measurement conditions	Approximate molecular weight (M_w, Daltons)
HYDROXYPROPYL CELLULOSE (HYDROXYPROPYL MS ~3.4–4.1)			
1500–3000	1	Brookfield (30 rpm/spindle 3)	1,150,000
4000–6500	2	Brookfield (60 rpm/spindle 4)	850,000
150–400	2	Brookfield (60 rpm/spindle 2)	370,000
150–400	5	Brookfield (60 rpm/spindle 2)	140,000
75–150	5	Brookfield (30 rpm/spindle 1)	95,000
300–600	10	Brookfield (30 rpm/spindle 2)	80,000
HYDROXYETHYL CELLULOSE (HYDROXYETHYL MS ~2.5)			
3500–5500	1	Brookfield (30 rpm/spindle 4)	1,300,000
1000–2500	1	Brookfield (30 rpm/spindle 3)	1,000,000
4500–6500	2	Brookfield (60 rpm/spindle 4)	720,000
250–400	2	Brookfield (60 rpm/spindle 2)	300,000
75–150	5	Brookfield (30 rpm/spindle 1)	90,000
HYDROXYPROPYL METHYLCELLULOSE (K TYPE (19–24 WT.% METHOXYL), E TYPE (28–30 WT.% METHOXYL), 7–12 WT.% HYDROXYPROPYL)			
160,000–240,000	2	Ubbelohde, too high for reliable measurement	1,200,000
80,000–120,000	2	Ubbelohde	1,000,000
19,200–36,000	2	Ubbelohde	675,000
11,250–21,000	2	Ubbelohde	575,000
3000–5600	2	Ubbelohde	400,000
SODIUM CARBOXYMETHYL CELLULOSE (NOMINAL CARBOXYMETHYL DS LEVELS OF 0.7, 0.9, 1.2)			
1500–3000	1	Brookfield (30 rpm/spindle 3)	725,000
1500–3100	2	Brookfield (30 rpm/spindle 3)	395,000
400–800	2	Brookfield (30 rpm/spindle 2)	250,000
25–50	2	Brookfield (60 rpm/spindle 1)	90,500
50–200	4	Brookfield (60 rpm/spindle 2)	49,000
ETHYL CELLULOSE (STANDARD OR N TYPE (48.0–49.5 WT.% ETHOXYL), T TYPE (49.6–51.0 WT.% ETHOXYL))[A]			
80–105	5	Ubbelohde 2C	215,000
40–52	5	Ubbelohde 2	160,000
18–24	5	Ubbelohde 2	140,000
12–16	5	Ubbelohde 1B	120,000
8–11	5	Ubbelohde 1B	75,000
6–8	5	Ubbelohde 1C	65,000

[a] *Estimated, based on viscosity data.*

derivatives, depending on type and viscosity grade, this ratio can range from 3 to values of 20 or so.

Typical examples of commonly used cellulose ethers and esters include the following[65]:

7.3.1.1 Hydroxypropyl cellulose

HPC is an ether of cellulose where some of the hydroxyl groups on the cellulose backbone have been hydroxypropylated (Fig. 7.43).

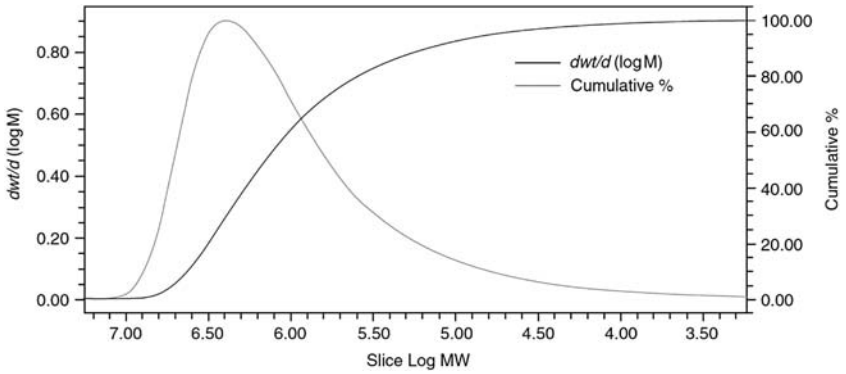


FIGURE 7.42 Representative molecular weight distribution profile of a high-viscosity grade (1500–3000 cP, 1% solution in water) sample of HPC. The molecular weight scale calibration in this example refers to narrow distribution polyethylene oxide/polyethylene glycol. Note that in this example, the net range of material spans roughly 2000–10,000,000 Da from the start to the end of the complete molecular weight distribution. This is a fairly typical situation for high-molecular-weight polymers.

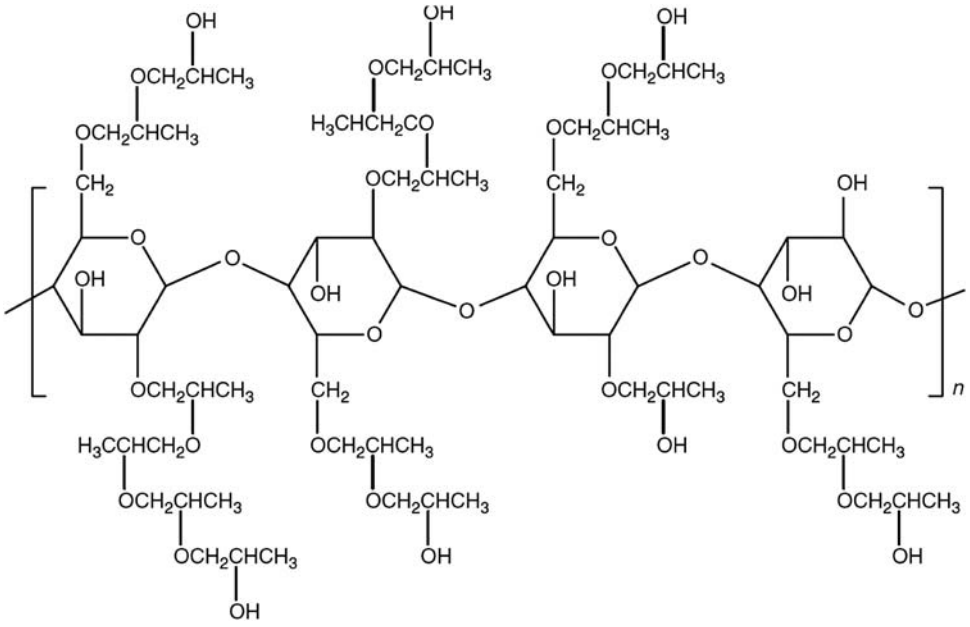


FIGURE 7.43 Typical structure for hydroxypropyl cellulose. The net molar substitution in this example is 4.0. Note that there is a distribution in the hydroxypropyl chain length for the pendant groups added to the base cellulose polymer. The relatively high hydroxypropyl content yields a somewhat hydrophobic thermoplastic polymer, which is sparingly soluble in hot water.

As mentioned earlier, because each of the added hydroxypropyl group introduces a secondary hydroxyl group, this can also be etherified during the preparation of HPC, giving rise to additional chain extension. When this occurs, the number of moles of hydroxypropyl groups per anhydroglucose ring, the molar substitution MS, will be higher than the degree of substitution DS (Fig. 7.43). To overcome the crystallinity of cellulose, HPC usually needs to have a MS about 4 in order to produce good solubility in water. Due to the high level of hydroxypropylation (around 70%), HPC is relatively hydrophobic, so it exhibits a LCST at 45°C. At temperatures below the LCST, HPC is readily soluble in water; however, above the LCST, HPC is not soluble. HPC is widely used in solid oral dosage forms as a binder, a film coating, controlled-release matrix former, and as extrusion aid. HPC is commercially available in six different viscosity grades corresponding to average molecular weights from 80,000 to 1,150,000 Da (Table 7.9).

7.3.1.2 *Hydroxypropyl methylcellulose*

Hydroxypropyl methylcellulose (HPMC or hypromellose) is a partly *O*-methylated and *O*-(2-hydroxypropylated) cellulose ether derivative (Fig. 7.44).

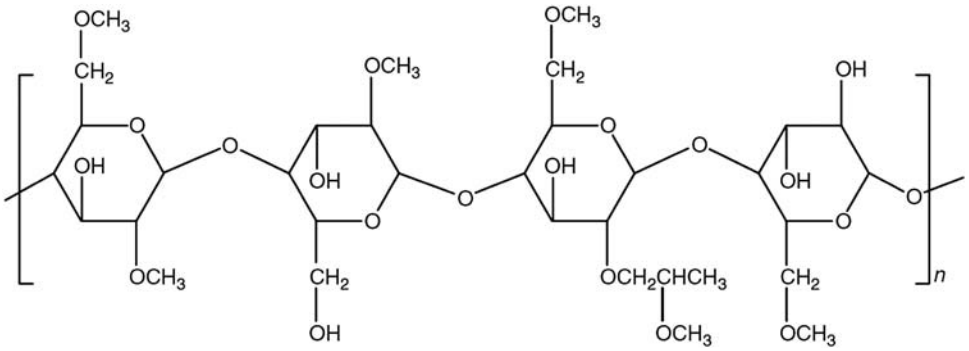


FIGURE 7.44 Typical structure for hydroxypropyl methylcellulose. This is an example of a mixed derivatized cellulosic containing both hydroxypropyl and methoxyl functionality. The specific system shown has a hydroxypropyl DS of 0.25 (~9.6 wt.%) and methoxyl DS of 1.5 (~23 wt.%). Due to the lower content of bound hydrophobic functionality relative to HPC shown in Fig. 7.38, this material will generally display better water solubility at higher temperatures.

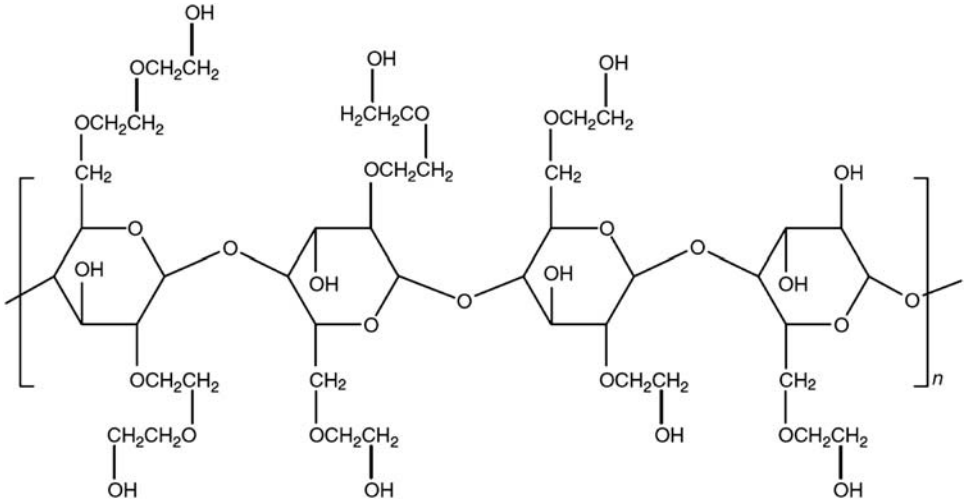


FIGURE 7.45 Typical structure for hydroxyethyl cellulose. As with hydroxypropyl cellulose, a distribution in hydroxyethyl oligomer chain lengths exists. In this example, the hydroxyethyl cellulose segment shown has a net molar substitution (MS) of 2.5, while the degree of substitution (DS) is 1.75 (=7/4).

HPMC is also widely used in solid oral dosage forms as a binder, a film coating, and a controlled-release matrix.[66,67] It is often the polymer of choice in the preparation of hydrophilic matrix tablets because of its rapid formation of a uniform, strong, and viscous gel layer, which protects the matrix from disintegration and controls the rate of drug release. It is commercially available in several types with different DS and MS; here the added hydroxypropyl group introduces a secondary hydroxyl group that can also be etherified during the preparation of HPMC, giving rise to additional chain extension. Currently, the USP/NF and other compendia provide definitions for different substitution types of HPMC using a four-digit number: for example, hypromellose 2208, hypromellose 2906, and hypromellose 2910, respectively, where the first two digits refer to the approximate percentage content of the more methoxy group (OCH_3), and the second two digits refer to the approximate percentage content of the more hydroxypropoxy group ($OCH_2CH(OH)CH_3$). Such variations in the ratios of methoxy and hydroxypropoxyl substitutions and molecular weight affect their properties such as organic solubility, thermal gelation temperature in aqueous solution, swelling, diffusion, and drug-release rate. In practice, hypromellose 2208 (also known as K chemistry or grade) and hypromellose 2910 (also known as E chemistry or grade) are the most widely used types of HPMC in modified-release formulations. These two types of HPMC are commercially available in many different viscosity grades corresponding to average molecular weights from 80,000 to 1,200,000 Da (Table 7.9).

7.3.1.3 *Hydroxyethyl cellulose*

HEC is a partially substituted polyhydroxyethyl ether of cellulose (Fig. 7.45).

It is used in ophthalmic and topical formulations as a viscosity-modifying agent and as a binder, controlled-release matrix former, and as a film-coating

agent in solid dosage forms. Similar to the synthesis of HPC, each of the added hydroxyethyl group introduces a secondary hydroxyl group that can be further etherified during the preparation of HEC, giving rise to an additional chain extension. In this case, the number of moles of hydroxyethyl groups per anhydroglucose ring, the molar substitution MS, will be higher than the degree of substitution DS. A typical structure of HEC with substitution sites on the anhydroglucose rings, giving rise to MS value of 2.5 and DS value of 1.75, is illustrated in Fig. 7.40. HEC dissolves quickly in water (hot or cold) to form a clear and smooth solution, which does not gel or precipitate even when heated to the boiling point of water. HEC is available commercially in several grades that vary in viscosity and DS, corresponding to average molecular weights from about 90,000 to 1,300,000 Da (Table 7.9).

7.3.1.4 Ethyl cellulose

Ethyl cellulose (EC) is a partly *O*-ethylated cellulose ether derivative (Fig. 7.46).

It is insoluble in water but soluble in a variety of solvents. EC is widely used in oral pharmaceutical formulations as a coating agent for tablets and granules to regulate the drug-release rate and to mask unpleasant taste, as well as in microencapsulation and in matrix tablets to achieve modified release. EC films can be applied from organic solvent systems, but this has largely been replaced by the use of a more environmentally friendly aqueous dispersion of EC without the need for organic solvents.[68–71] The substitution level or the ethoxyl content directly impacts the properties of the resulting EC. A typical structure of EC with a DS value of 2.5 corresponding to a 48.5 wt.% ethoxyl content is shown in Fig. 7.46. If the DS of EC is increased to 2.8, a largely insoluble material would result, underscoring the interplay of solid-state packing defects in modulating the solubility of cellulose derivatives. Currently, USP-NF and EP define the ethoxyl content of EC to be between 44.0 and 51.0 wt.%. EC is available commercially in several types that vary in viscosity and ethoxyl content, corresponding to average molecular weights from about 65,000 to 215,000 Da; the most widely used is the standard or *N* type that has 48.0–49.5 wt.% ethoxyl content (Table 7.9). EC can also be used as a binder/matrix former for matrix tablet formulations for modified-release applications.[72]

7.3.1.5 Methyl cellulose

Methyl cellulose (MC) is a methyl ester of cellulose (Fig. 7.47) that contains 27.5–31.5% of the methoxy groups. A typical structure of MC with DS value of

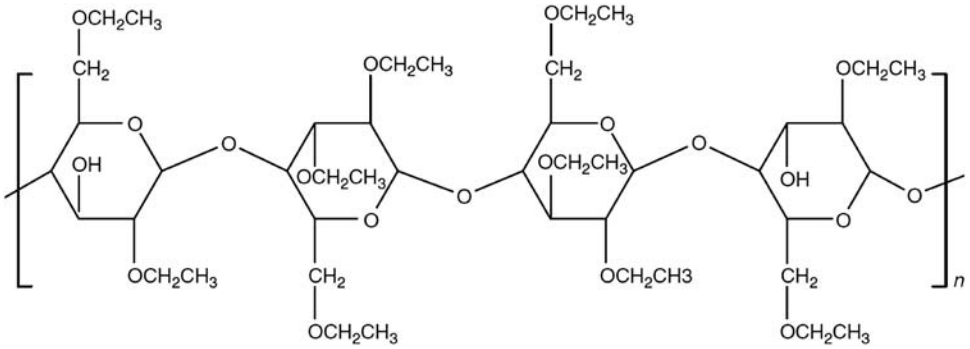


FIGURE 7.46 Typical structure for ethyl cellulose with a net degree of substitution of 2.5. This corresponds to a wt.% ethoxyl of 48.5, which is representative of an *N* or standard grade of this material. If the net DS of the ethyl cellulose is substantially increased to the order of 2.8, a largely insoluble material results, underscoring the interplay of solid-state packing defects in modulating cellulose derivative solubility.

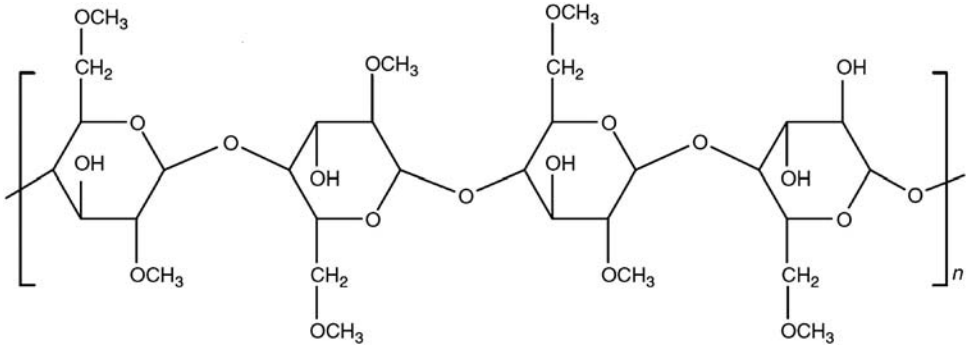


FIGURE 7.47 Typical structure for methyl cellulose with a net degree of substitution of 1.75, which corresponds to a wt.% methoxyl level of 29.1.

1.75 corresponding to 29.1 wt.% methoxyl content is shown in Fig. 7.47. MC is soluble in water, and its aqueous solution exhibits thermal gelation properties. A wide range of viscosity grades (5–75,000 cP at 2%) corresponding to average molecular weight range of 10,000–220,000 Da are available commercially. MC is widely used in oral solid pharmaceutical formulations as a binder, a coating agent, and a disintegrant for tablets, as well as in matrix tablets to achieve sustained release.

7.3.1.6 Sodium carboxymethyl cellulose

Sodium carboxymethyl cellulose (NaCMC) is the sodium salt of the carboxymethyl ether of cellulose, an anionic derivative (Fig. 7.5).

It is widely used in oral, ophthalmic, injectable, and topical pharmaceutical formulations, primarily for its viscosity-increasing and gelling properties. For solid dosage forms, it is mostly used as a binder or matrix former. Pharmaceutical grades of NaCMC are available commercially at DS values of 0.7 and 0.9, with a corresponding sodium content of 6.5–9.5 wt.% (USP/EP), and at a DS value of 1.2 with sodium content of 10.4–12 wt.% (USP). NaCMC is highly soluble in water at all temperatures, forming clear solutions with viscosities depending on the substitution grade and concentration. NaCMC is available in several different viscosity grades, corresponding to average molecular weights from about 49,000 to 725,000 Da (Table 7.9).

7.3.1.7 Cellulose acetate

Cellulose acetate (CA) is cellulose with its hydroxyl groups partially or completely acetylated (Fig. 7.48).

It is insoluble in water but soluble in a variety of organic solvents. CA is available in a wide range of acetyl contents (29.0–44.8%) and chain lengths, with molecular weights ranging from 30,000 to 60,000. CA is widely used in pharmaceutical formulations as semipermeable coatings on osmotic tablets to regulate the osmotic water influx, thereby controlling the drug-release rates.[73–76] With appropriate additives, it is also used as coatings on tablets or granules or as shell material for microcapsules to achieve controlled drug release, or as coatings on tablets or granules for taste-masking purposes. CA can also be employed to formulate matrix tablet prepared by direct compression.

7.3.1.8 Cellulose derivatives with pH-dependent solubility

HPMCAS (Fig. 7.49) is produced by the esterification of HPMC with acetic anhydride and succinic anhydride in a reaction medium of a carboxylic acid, such as acetic acid, and using an alkali carboxylate, such as sodium acetate, as catalyst.

HPMCAS retains the structure diversity of the starting material HPMC where the added hydroxypropyl group introduces a secondary hydroxyl group, which can also be etherified during the preparation of HPMC, giving rise to an additional chain extension. It is available in several grades with varying extent of substitution, mainly of acetyl (2.0–16.0%) and succinoyl (4.0–28.0%) groups with molecular weights ranging from 55,000 to 93,000 Da. HPMCAS is insoluble in gastric fluid but starts to swell and dissolve at a pH above 5 depending on the extent of substitution. The dissolution pH increases as the ratio of acetyl over succinoyl substitution increases. Traditionally, HPMCAS has been used in enteric film coating of tablets and multiparticulates.[77] For aqueous film-coating formulations, a dispersion of HPMCAS fine powder and triethyl citrate (as a plasticizer) in water is commonly utilized. Organic solvents can also be used as vehicles for applying this polymer as a film coating. However, today HPMCAS is the most widely utilized amorphous solid dispersion polymer for drug solubilization. It is most easily used via spray drying, but hot-melt extrusion is also becoming increasingly common.[78,79]

Cellulose acetate phthalate (CAP) (Fig. 7.50) is a product of esterification of cellulose acetate (CA) with phthalic anhydride in the presence of a tertiary organic base such as pyridine or a strong acid such as sulfuric acid containing 21.5–26.0% of acetyl groups

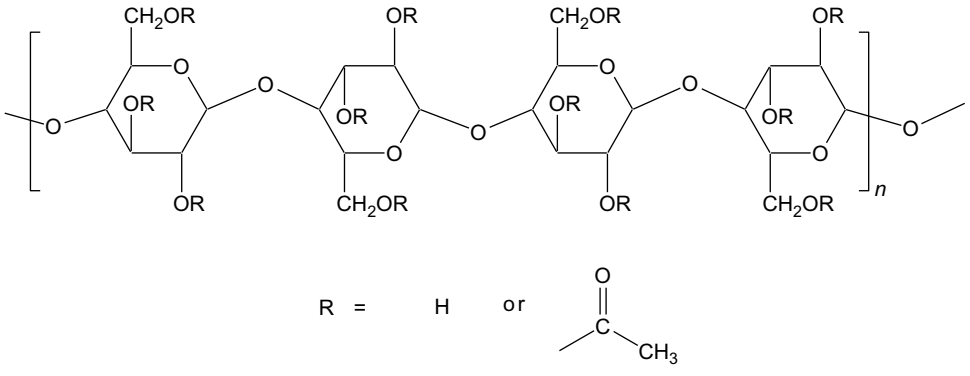


FIGURE 7.48 Typical structure for cellulose acetate (CA).

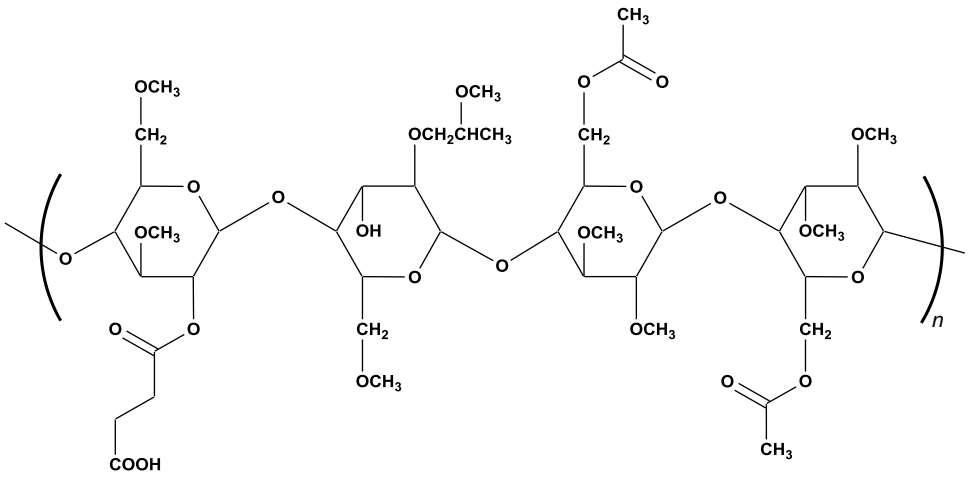


FIGURE 7.49 Typical structure for hydroxypropyl methylcellulose acetate succinate (HPMCAS).

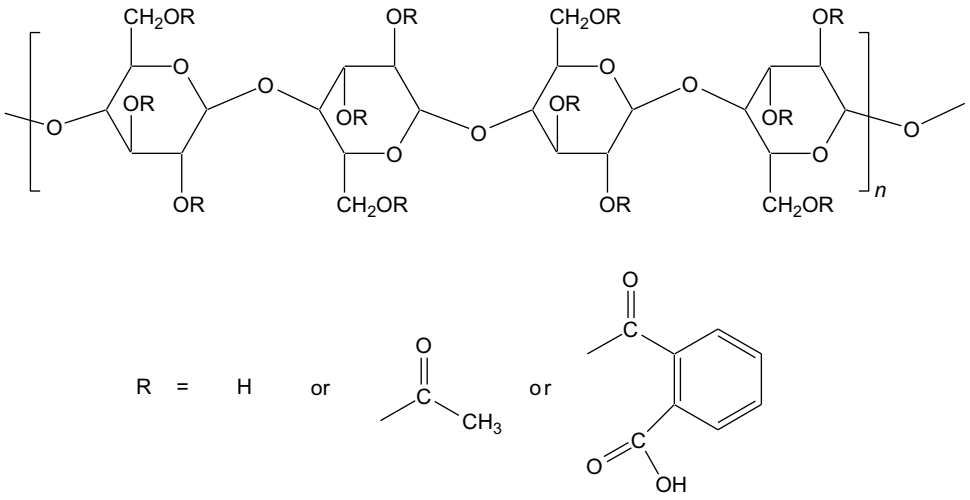


FIGURE 7.50 Typical structure for cellulose acetate phthalate (CAP).

and 30.0–36.0% of phthalyl groups with a viscosity (15% in acetone) ranging from 50 to 90 cP.

CAP has pH-dependent solubility and is soluble in aqueous solutions at a pH above 6. CAP is mainly used as an enteric film-coating material or as a matrix binder for tablets and capsules. It provides protection of API in the strongly acidic gastric fluid, but it dissolves in mildly acidic or neutral intestinal environment to make the drug available for absorption. CAP is commonly applied to solid-dosage forms either by coating from an aqueous dispersion or a solution of organic solvent or by direct compression.

7.3.2 Synthetic polymers

7.3.2.1 Acrylic acid polymers

Synthetic polymers based on acrylic, especially methacrylic acids, have found extensive applications in the oral solid formulations to protect the active ingredients as versatile film former for coatings or to release the drug in a controlled manner as a matrix former or a diffusion barrier for a reservoir system.

Such polymers can be made as cationic, anionic, and neutral (nonionic) polymers.

7.3.2.1.1 Polyacrylic acid (carbomer; carbopol)

Carbomers (Fig. 7.51) are synthetic high-molecular-weight polyacrylic acids cross-linked with allyl sucrose or allyl pentaerythritol and contain between 56 and 68% w/w carboxylic acid groups. The molecular weight of carbomer is estimated to be at 7×10^5 to 4×10^9 Da. As three-dimensionally cross-linked microgels, carbomers do not dissolve but swell to a remarkable extent in water after neutralization to form a gel.[80] Carbomers are used in liquid or semisolid pharmaceutical formulations as rheology modifiers. Carbomers having low residual solvent content, such as Carbopol 971P NF or

Carbopol 974P NF, may be used in oral solid dosage forms as binders or matrix tablet formers.[81–84]

7.3.2.1.2 Polymethacrylate

Polymethacrylates (Fig. 7.52) are synthetic linear copolymers prepared by free-radical polymerization. They may exist as cationic, anionic, and neutral (nonionic) polymers depending on the starting monomers for preparing the polymers. The functionality and applications of such polymers depend on the structures and ionic charges of the polymers.[85–88]

Anionic polymers: Anionic polymethacrylate polymers contain methacrylic acid functional groups, which dissociate and render the polymer soluble at the higher pH of the small intestine and colon. Anionic polymers containing methacrylic acid–ethyl acrylate copolymer (1:1) in coating products (eg, Eudragit® L 30-D or Kollicoat® MAE 30 DP) are soluble from a pH 5.5 but insoluble at the low pH in the stomach. Such polymers offer enteric protection to many active ingredients in oral solid formulations in the gastric environment and trigger drug release at a selected pH for targeted drug delivery in the GI tract. Anionic copolymers of methacrylic acid and methyl methacrylate at a ratio of approximately 1:1 (Type A; eg, Eudragit L) and approximately 1:2 (Type B; eg, in Eudragit S). Both polymers are readily soluble in neutral to weakly alkaline conditions (from pH 6 for

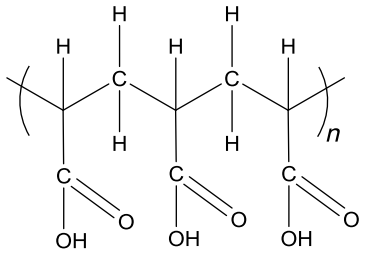
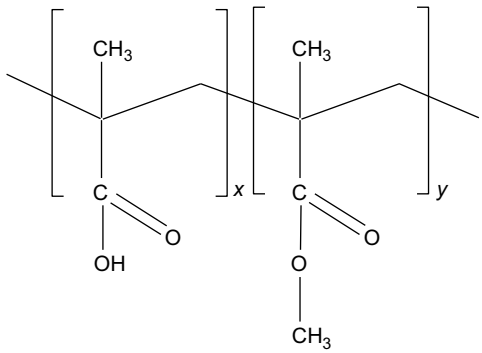
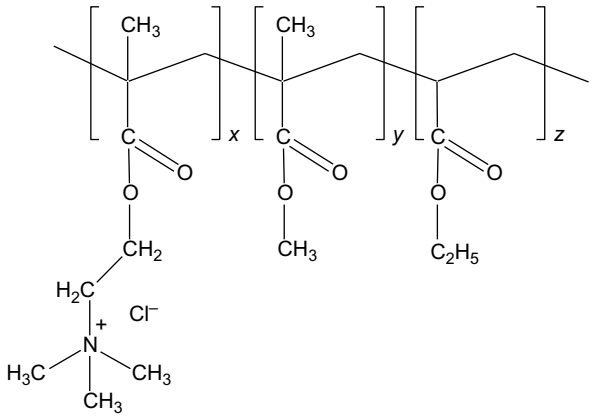


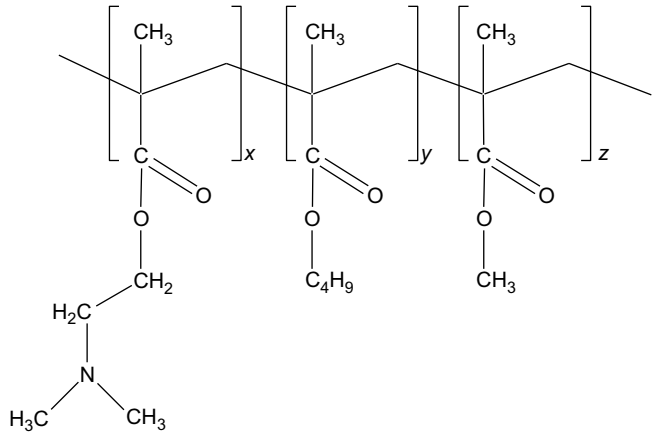
FIGURE 7.51 Typical structure for carbomers.



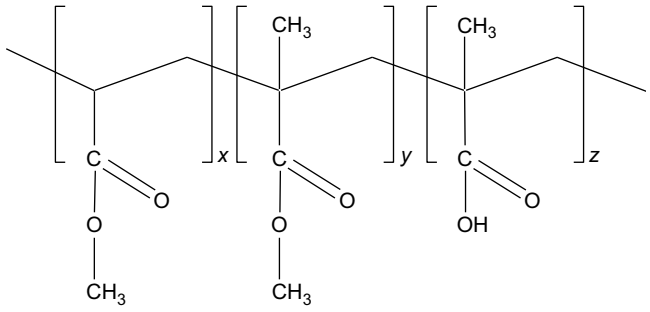
Methacrylic acid—methyl methacrylate copolymer (1:1 or 1:2)



Poly(ethylacrylate-co-methyl methacrylate-co-trimethylammonioethyl methacrylate chloride) (1:2:0.2 or 1:2:0.1)



Poly(butyl methacrylate-co-(2-dimethylaminoethyl) methacrylate-co-methyl methacrylate) (1:2:1)



Poly(methyl acrylate-co-methyl methacrylate-co-methacrylic acid) (7:3:1)

FIGURE 7.52 Typical structure for polymethacrylates.

Type A and from pH 7 for Type B); the latter enables drug delivery to the colon, since the polymers becomes soluble at pH >7.0.

Cationic polymers: Cationic polymethacrylate polymers typically consist of copolymers of ethyl acrylate, methyl methacrylate, and a low content of methacrylic acid ester with quaternary ammonium groups. The ammonium groups are present as salts, which increase the polymer swelling in aqueous media, thereby making the polymers more permeable. Cationic polymethacrylate containing more quaternary ammonium groups (eg, Eudragit RL 100) is, therefore, more permeable than the one containing less quaternary ammonium groups (eg, Eudragit RS 100) as a consequence. Both polymethacrylate polymers are commercially available as a fine powder (eg, Eudragit RS PO), a solution in organic solvents (eg, Eudragit RS 12.5), or an aqueous dispersion (eg, Eudragit RS 30 D). Both polymethacrylate polymers can be mixed in any ratio to modulate the drug-release rate in addition to well-known strategies of varying coating thickness and incorporation of pore formers.

Another available cationic polymethacrylate is a copolymer based on dimethylaminoethyl methacrylate, butyl methacrylate, and methyl methacrylate (eg, Eudragit E 100). It is soluble in gastric fluid below pH 5 but becomes swellable and permeable, but not soluble, above pH 5. This cationic polymethacrylate can be used for taste-masking applications.

Neutral polymers: Neutral polymethacrylates based on copolymers of ethyl acrylate and methyl methacrylate (2:1) do not contain ionic groups and, therefore, only swell in aqueous media independently of pH without dissolving. These are generally available as aqueous dispersions with a solid content at 30% and 40% respectively (eg, Eudragit NE 30 D and Eudragit NE 40 D). Films prepared from these are insoluble in water but will swell and become permeable when they are in contact with water and exhibit pH-independent drug permeability. Such neutral polymethacrylate aqueous dispersions can be used for controlled-release coatings and wet-granulation binders.

7.3.2.2 Polyvinylpyrrolidone

7.3.2.2.1 Povidone

Povidone (Fig. 7.53) is a linear polymer of 1-vinyl-2-pyrrolidinone monomer with different degrees of polymerization, which results in polymers of a wide range of molecular weights (2,500−3,000,000 Da).

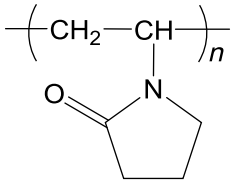


FIGURE 7.53 Chemical structure for polyvinylpyrrolidone (PVP).

Povidone is water soluble, with the maximum concentration being limited only by the solution viscosity. Being nonionic, the viscosity of povidone in an aqueous solution is unaffected by the pH or salt concentration. Povidone is widely used in solid-dosage forms. In tablet formulations, povidone solution is used as wet-granulation binder. Povidone can also be incorporated to powder blends in the dry form and granulated in situ by the addition of water, alcohol, or hydroalcoholic solutions. Povidone may be used as a dissolution-enhancing agent in oral formulations for poorly soluble drugs involving processes such as spray drying and hot-melt extrusion.[79,89] Povidone is a well-known complexing agent; for example, the ability to complex with iodine is of particular commercial importance, providing an antibacterial povidone iodine solution and powder that are widely used in medical and hospital settings. Povidone solutions may also be used as a film former for coating formulations or as binders for drug layering onto a substrate such as sugar or microcrystalline cellulose beads. Povidone is additionally used as a suspending, stabilizing, or viscosity-increasing agent in a number of oral suspensions and solutions. Increasingly due to its advantageous solubility in PEG 400 and glycerin, povidone is also used as a viscosifying and stabilizing agent in liquid-filled gelatin capsules.

7.3.2.2.2 Crospovidone

The preparation of vinylpyrrolidone for crospovidone is similar to povidone. The vinylpyrrolidone is polymerized in solution using a catalyst to produce crospovidone by a popcorn polymerization process. Crospovidone is a white, free-flowing, practically tasteless and odorless, hygroscopic powder.

Crospovidone is a cross-linked, water-insoluble superdisintegrant. It is usually incorporated dry in the running powder of a tablet formulation, but it can also be processed by wet and dry granulation methods. Typical use levels in tablet formulations range from 2% to 5%, but in orally disintegrating tablets in particular and some capsule formulations, crospovidone is used at much higher levels up to 30%. Similar to sodium starch glycolate, crospovidone disintegrates tablets mainly by swelling, with little tendency to form gels. Crospovidone can also be used for solubility enhancement of poorly soluble drugs in the process of coevaporation. This process enables the drug adsorption onto crospovidone in the presence of a suitable solvent, and the solvent is then evaporated to provide a solid mixture with a faster drug dissolution rate.

7.3.2.3 *Polyvinyl alcohol (PVA)*

Polyvinyl alcohol (PVA) is produced through the hydrolysis of polyvinyl acetate. The repeating unit of vinyl alcohol is not used as the starting material because it is unstable and cannot be isolated. The hydrolysis of polyvinyl acetate proceeds rapidly in methanol, ethanol, or a mixture of alcohol and methyl acetate, using alkalis or mineral acids as catalysts.

PVA occurs as an odorless, white to cream-colored granular powder. It is a linear water-soluble polymer represented by the formula $(C_2H_4O)_n$ (see structure in Fig. 7.54). The value of n for commercially available PVA is between 500 and 5000, equivalent to a molecular weight range of 20,000−200,000. Various grades of PVA are commercially available. The DP and the degree of hydrolysis are the two key factors affecting their physical properties. Typically, pharmaceutical grades are partially hydrolyzed materials.[65]

The predominant commercial use of polyvinyl alcohol in solid oral dosage forms is as an immediate-release, film-coating polymer. It is also used as a viscosity-enhancing agent in ophthalmic and topical formulations.

7.3.2.4 *Polyethylene oxide (PEO) and polyethylene glycol (PEG)*

Polyethylene oxide (PEO) and polyethylene glycol (PEG) have the same CAS registry number 25322-68. Both PEO and PEG are nonionic homopolymers of ethylene oxide, sharing the same formula $(CH_2CH_2O)_n$ (Fig. 7.55) where n represents the average number of oxyethylene groups, with $n = 5-182$ for typical PEGs (avg. MW 200-8000) and much larger n's for PEOs (avg. MW 100,000 up to several million).

7.3.2.4.1 Polyethylene glycol (PEG)

Polyethylene glycol (PEG) is formed by the reaction of ethylene oxide and water under pressure in the presence of a catalyst.

PEG is soluble in water and miscible in all proportions with any other grade of PEG (after melting, if necessary). Liquid PEGs are soluble in acetone, alcohols, benzene, glycerin, and glycols. Solid PEGs are soluble in acetone, dichloromethane, ethanol (95%), and methanol; they are slightly soluble in aliphatic hydrocarbons and ether but insoluble in fats, fixed oils, and mineral oil. Aqueous solutions of higher-molecular-weight grades may form gels.

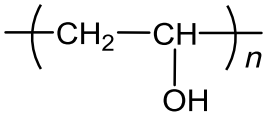


FIGURE 7.54 Chemical structure for polyvinyl alcohol (PVA).

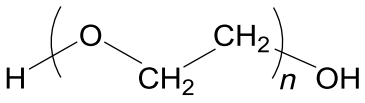


FIGURE 7.55 Chemical structure for polyethylene glycol (PEG) and polyethylene oxide (PEO).

In tablet formulations, PEGs of a high molecular weight can enhance the effectiveness of tablet binders and impart plasticity to granules when it is used in conjunction with another binder. PEG may prolong disintegration if used above 5% w/w. Polyethylene glycol grades with molecular weights of 6000 and above can be used as lubricants, particularly for soluble tablets. The lubricant action is not as good as that of magnesium stearate, and stickiness may develop if the material becomes too warm during compression.

In coating formulations, low-molecular-weight PEG is mainly used as a plasticizer. Higher molecular weights can be useful as hydrophilic polishing materials. Solid PEGs are also used to enhance the ductility of the coating membrane and avoid rupture of the coating film when the coated microcapsules are compressed into tablets. The presence of PEGs in film coats, especially of liquid grades, tends to increase their water permeability and may reduce protection against low pH in enteric-coating films.[65]

PEGs may also be used to enhance the aqueous solubility of poorly soluble drugs by making solid dispersions with an appropriate grade of polyethylene glycol by either spray drying or hot-melt extrusion.

Polyethylene glycols have been used in the preparation of urethane hydrogels, which are used as controlled-release agents. Polyethylene glycol has also been used in insulin-loaded microparticles for the oral delivery of insulin; it has been used in inhalation preparations to improve aerosolization.

7.3.2.4.2 Polyethylene oxide (PEO)

Polyethylene oxide (PEO) is prepared by the polymerization of ethylene oxide with the use of a suitable catalyst.

PEO is soluble in water and some commonly used organic solvents such as acetonitrile, chloroform, and methylene chloride. It is insoluble in aliphatic hydrocarbons, ethylene glycol, and most alcohols. It may contain up to 3% of silicon dioxide or suitable antioxidant.

Higher-molecular-weight PEO is predominantly used in extended-release formulations as a hydrophilic matrix former at levels ranging from 5% to 75%. PEO can be used as a tablet-and-extrudate forming material or an aid in hot-melt extrusion. The relationship between the swelling capacity and the molecular

weight is a good guide when selecting products for use in immediate- or sustained-release matrix formulations. Polyethylene oxide has also been shown to have mucoadhesive properties. PEO films exhibit good lubricity under wet conditions.[65]

PEOs are an effective viscosity enhancer at low-use level, although alcohol is usually added to water-based formulations to provide improved viscosity stability.

7.3.2.5 *Ion-exchange resins*

Ion-exchange resins are generally made from methacrylic acid, sulfonated styrene, and divinylbenzene (DVB). A generic structure of such cation-exchange resin is shown in Fig. 7.56.

Cation exchangers are anionic polymers that contain carboxyl or sulfate groups with hydrogen, potassium, and sodium as counterions. Cation exchangers with weak acidity are made from a polymer of methacrylic acid (containing COOH) cross-linked by DVB. The counterion of the acidic carboxyl group is either hydrogen (as in polacrilex resin) or potassium (as in polacrilin potassium). To make a cation-exchange resin with a stronger acidity, water-insoluble styrene is used to prepare the polymer, which is sulfonated to make it hydrophilic. DVB is also used to cross-link the polymer, and the counterion of the sulfate group (SO_3) is generally sodium.

Ion-exchange resins swell in an aqueous medium. The DVB–cross-linked potassium methacrylate copolymer possesses such a high swelling capacity that it is used as a disintegrant for oral solid formulations.

Ion-exchange resins have fixed ionic functional groups that can provide binding of ionic drugs. Release of the bound drugs requires an exchange with counterions such as hydrogen or sodium, which are

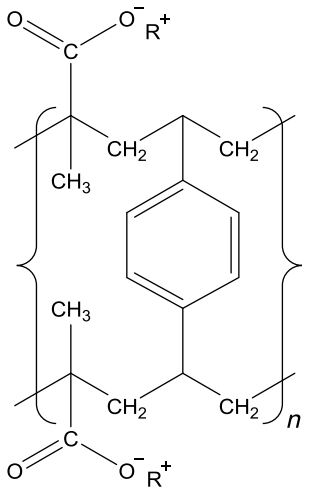


FIGURE 7.56 Generic structure of a cation-exchange resin based on polymethacrylic acid, where R = hydrogen for polacrilex resin, R = potassium (partial) for polacrilin potassium.

available in the gastrointestinal tract. Because of their unique properties, the ion-exchange resins are generally used for taste masking, drug stabilization, and sustained-release applications or a zero-order release due to their high swelling capacity.[90,91]

The commercial product of sodium polystyrene sulfonate (Amberlite™ IRP69) is used to treat hyperkalemia. On the other hand, cationic ion-exchange resins with the ability to exchange anions carry quaternary ammonium groups, $-N^+(R)_3$ with chlorine as a counterion. Cholestyramine resin (Duolite™ AP143) is cationic styrene DVB polymer that is an anion exchanger and is used to reduce cholesterol or to sequestrate the bile acid.

7.4 CONCLUSION

In many respects, polymeric materials have a preeminent position among excipients, as they form the backbone of modern pharmaceutical practice and drug-delivery technologies in particular. With the exception of a few technologies, the majority of bioavailability enhancement, modified-release, and controlled-release technologies rely on the unique properties of polymers to achieve drug-release control through a variety of mechanisms. These may include dissolved drug diffusion through water-insoluble, polymeric film coatings or diffusion and erosion from water-soluble, hydrogel-forming polymeric systems. Some of the obvious, inherent properties of polymers that continue to enable the development of new technologies include their large molecular weight, which facilitates entanglement and network formation, thus allowing diffusion control. Also important is the availability of large numbers of different polymers, which are generally regarded as safe and which have good processing and manufacturing properties such as solubility in aqueous systems and common nonaqueous solvents, good stability, good plastic deformation and compaction properties, and excellent film-forming properties.

As highlighted in this chapter, the key to understanding and productively using polymeric materials in solid dosage form design and general industrial practice is to appreciate their unique properties as compared to well-defined, discrete small molecules. In particular, unlike small molecules with a well-defined and discrete structure, state, and properties, polymers are best viewed in terms of a continuum and distribution of structural and physical property characteristics and states. Furthermore, while structure–property relationships are frequently complex and multifactorial, we hope to have conveyed that even with the application of only a small set of relatively simple rules and principles, it is possible to make

more rational and science-based decisions with regard to polymer selection to achieve the desired effects and robust systems. This is particularly important in the context of the current industry focus on QbD and the fact that polymeric materials frequently comprise a significantly greater proportion of the modern dosage form than does the drug. Looking to the future, it is clear that polymers will continue to be a tool in the development of solid oral dosage forms.

References

1. Mark J, Ngai K, Graessley W, Mandelkern L, Samulski E, Koenig J, et al. *Physical properties of polymers*. 3rd ed. Cambridge: Cambridge University Press; 2004.
2. Sperling LH. *Introduction to physical polymer science*. 4th ed. New York, NY: Wiley-Interscience; 2005.
3. Ferry JD. *Viscoelastic properties of polymers*. 3rd ed. New York, NY: Wiley-Interscience; 1980.
4. Odian G. *Principles of polymerization*. 4th ed. New York, NY: Wiley-Interscience; 2004.
5. Tanford C. *Physical chemistry of macromolecules*. New York, NY: John Wiley & Sons; 1961.
6. Flory PJ. *Principles of polymer chemistry*. Ithaca, NY: Cornell University Press; 1953.
7. Landoll LM. Nonionic polymer surfactants. *J Polymer Sci, Polymer Chem* 1982;**20**:443–55.
8. Casas JA, García-Ochoa F. Viscosity of solutions of xanthan/locust bean gum mixtures. *J Sci Food Agri* 1999;**79**:25–31.
9. Momcilovic D, Schagerlof H, Rome D, Jornten-Karlsson M, Karlsson K-E, Wittgren B, et al. Derivatization using dimethylamine for tandem mass spectrometric structure analysis of enzymatically and acidically depolymerized methyl cellulose. *Analyt Chem* 2005;**77**:2948–59.
10. Richardson S, Andersson T, Brinkmalm G, Wittgren B. Analytical approaches to improved characterization of substitution in hydroxypropyl cellulose. *Analyt Chem* 2003;**75**:6077–83.
11. Schagerlof H, Richardson S, Momcilovic D, Brinkmalm G, Wittgren B, Tjerneld F. Characterization of chemical substitution of hydroxypropyl cellulose using enzymatic degradation. *Biomacromolecules* 2006;**7**:80–5.
12. Schagerlof H, Johansson M, Richardson S, Brinkmalm G, Wittgren B, Tjerneld F. Substituent distribution and clouding behavior of hydroxypropyl methyl cellulose analyzed using enzymatic degradation. *Biomacromolecules* 2006;**7**:3474–81.
13. Wittgren, B. Wood material research—importance for industry. In: Presentation at the Opening Seminar, Wood Material Science Research Programme; Tuesday, April 6, 2004, Marina Congress Center. Katajanokanlaituri 6, Helsinki, Finland.
14. Barth HG, Mays JW, editors. *Modern methods of polymer characterization*. New York, NY: Wiley-Interscience; 1991.
15. Cooper AR. *Determination of molecular weight*. 1st ed. New York, NY: Wiley-Interscience; 1989.
16. Mori S, Barth HG. *Size exclusion chromatography*. New York, NY: Springer-Verlag; 1999.
17. Macosko C. *Rheology, principles, measurements and applications*. New York, NY: VCH; 1994.
18. Barnes HA, Hutton JF, Walters K. *An introduction to rheology*. Amsterdam: Elsevier; 1989.
19. Lapasin R, Pricl S. *Rheology of industrial polysaccharides*. London: Chapman and Hall; 1995.
20. Morrison FA. *Understanding rheology*. Oxford: Oxford University Press; 2001.
21. Dealy JM. Official nomenclature for material functions describing the response of a viscoelastic fluid to various shearing and extensional deformations. *J Rheol* 1995;**39**:253–65.
22. Sanchez IC, Lacombe RH. An elementary molecular theory of classical fluids: Pure fluids. *J Phys Chem* 1976;**80**:2352–62.
23. Lacombe RH, Sanchez IC. Statistical thermodynamics of fluid mixtures. *J Phys Chem* 1976;**80**:2568–80.
24. Sanchez IC, Lacombe RH. Statistical thermodynamics of polymer solutions. *Macromolecules* 1978;**11**:1145–56.
25. Hildebrand JH, Prausnitz JM, Scott RL. *Regular and related solutions*. New York, NY: Van Nostrand Reinhold; 1970.
26. Hansen CM. The universality of the solubility parameter concept. *I & E C Product Res Dev* 1969;**8**:2–11.
27. Barton AFM. *CRC handbook of solubility parameters and other cohesion parameters*. 4th ed. Boca Raton, FL: CRC Press; 1988.
28. Archer WL. Determination of Hansen solubility parameters for selected cellulose ether derivatives. *Ind Eng Chem Res* 1991;**30**:2292–8.
29. Klucel® Hydroxypropylcellulose: Physical and Chemical Properties. Ashland Specialty Ingredients, Ashland, Wilmington DE (250–2F, rev. 10–01 500).
30. Erhenstein GW, Theriault RP. *Polymeric materials: structure, properties, applications*. Munich: Hanser-Verlag; 2001.
31. Staudinger H, Heuer W. Highly polymerized compounds. XXXIII. A relationship between the viscosity and the molecular weight of polystyrenes (translated). *Ber Bunsenges Phys Chem* 1930;**63B**:222–34.
32. Billmeyer FW. *Textbook of polymer science*. 3rd ed. New York, NY: John Wiley & Sons; 1984. p. 242–357.
33. Dürig, T. et al. Fundamental studies on modified release pellets coated with ethylcellulose. Ashland Pharmaceutical Technology Report, PTR-33; 2005.
34. Hjartstam J, Borg K, Lindstedt B. The effect of tensile stress on the permeability of free films of ethylcellulose containing hydroxypropylmethyl cellulose. *Int. J Pharm* 1990;**61**:101–7.
35. Rowe RC. The effect of the molecular weight of ethyl cellulose on the drug release properties of mixed films of ethyl cellulose and hydroxypropyl methylcellulose. *Int J Pharm* 1986;**29**:37–41.
36. Prodduturi S, Manek V, Kolling WM, Stodghill SP, Repka MA. Water vapor sorption of hot melt extruded hydroxypropyl cellulose films: Effect on physico-mechanical properties, release characteristics and stability. *J Pharm Sci* 2004;**93**:3047–56.
37. Skinner, G.W., Harcum, W.W., Lusvardi, K.M., Lau, S.-F. & Dürig, T. Evaluation of hydroxypropylcellulose as a direct compression binder. Ashland Pharmaceutical Technology Report, PTR-25; 2003.
38. Picker-Freyer KM, Dürig T. Physical mechanical and tablet formation properties of hydroxypropylcellulose: In pure form and in mixtures. *AAPS PharmSciTech* 2007;**8** Article 92, <http://www.aapspharmscitech.org>.
39. Katikaneni PR, Upadrashta SM, Rowlings CE, Neau SH, Hileman GA. Consolidation of ethylcellulose: effect of particle size, press speed and lubricants. *Int J Pharm* 1995;**117**:13–21.
40. Fox TG, Flory PJ. Second-order transition temperatures and related properties of polystyrene. I, Influence of polymer molecular weight. *J Appl Phys* 1950;**21**:581–91.
41. Ju RTC, Nixon PR, Patel MV. Drug release from hydrophilic matrices. I. New scaling laws predicting polymer and drug release based on the polymer disentanglement concentration and the diffusion layer. *J Pharm Sci* 1995;**84**:1455–63.
42. Reynolds TD, Gehrke SH, Hussain AS, Shenouda LS. Polymer erosion and drug release characterization of hydroxypropyl methylcellulose matrices. *J Pharm Sci* 1998;**87**:1115–23.

43. Klug ED. Some properties of water soluble hydroxyalkyl celluloses and their derivatives. *J Polym Sci Part C* 1971;**36**:491–508.
44. Baumgartner S, Smid-Korbar J, Vrecer F, Kristl J. Physical and technological parameters influencing floating properties of matrix tablets based on cellulose ethers. *STP Pharm Sci* 1998;**8**:285–90.
45. Bi, V., Lester, J., Tewari, D. & Dürig, T. A rational approach for selection of HPMCAS grade in solid dispersion formulations. AAPS Poster W5274; 2013. www.ashland.com/Ashland/Static/Documents/ASI/Pharmaceutical/AAPS_2013_HPMCAS-Grade%20Selection_Final.pdf.
46. Tanno F, Nishiyam Y, Kobubo H, Obara S. Evaluation of hypromellose acetate succinate (HPMCAS) as a carrier in solid dispersions. *Drug Dev Ind Pharm* 2004;**30**:9–17.
47. Rajabi-Siahboomi AR, Nokhodchi A. Compression properties of methylcellulose and hydroxypropylmethylcellulose polymers. *Pharm Pharmacol Commun* 1999;**5**:67–71.
48. Wood LA. Glass transition temperatures of copolymers. *J Polym Sci* 1958;**28**:319–30.
49. BeMiller JN, Whistler RL, editors. *Industrial gums—polysaccharides and their derivatives*. 3rd ed. New York, NY: Academic Press; 1993.
50. Ott, E., Spurlin, H.M. & Grafflin, M.W. (eds Parts I-III), Bikales, N. M & Segal, L. (eds Parts IV, V). (1954). Cellulose and Cellulose Derivatives, Vol. V, High Polymers. Wiley Interscience, New York, NY.
51. Kimura M, Shimada Y, Oshima T, Shibata D, Hattori N, Yonezawa Y, et al. The evaluation of powdered cellulose as a pharmaceutical excipient. *J Pharm Sci Tech Japan* 2002;**62**:113–23.
52. Kothari SH, et al. Comparative evaluations of powder and mechanical properties of low crystallinity celluloses, microcrystalline celluloses, and powdered celluloses. *Int J Pharm* 2002;**232**:69–80.
53. Levis SR, Deasy PB. Production and evaluation of size-reduced grades of microcrystalline cellulose. *Int J Pharm* 2001;**213**:13–24.
54. Sun CC. Mechanism of moisture induced variations in true density and compaction properties of microcrystalline cellulose. *Int J Pharm* 2008;**346**:93–101.
55. Crouter A, Briens L. The Effect of Moisture on the Flowability of Pharmaceutical Excipients. *AAPS PharmSciTech* 2014;**15**:65–74.
56. Lerk CF, Bolhuis GK. Comparative evaluation of excipients for direct compression I. *Pharm Weekbl* 1973;**108**:469–81.
57. Lerk CF, et al. Comparative evaluation of excipients for direct compression II. *Pharm Weekbl* 1974;**109**:945–55.
58. Doelker E, et al. Comparative tableting properties of sixteen microcrystalline celluloses. *Drug Dev Ind Pharm* 1987;**13**:1847–75.
59. Doelker E. Comparative compaction properties of various microcrystalline cellulose types and generic products. *Drug Dev Ind Pharm* 1993;**19**:2399–471.
60. Thoorens G, Krier F, Leclercq B, Carlin B, Evrard B. Microcrystalline cellulose, a direct compression binder in a quality by design environment—a review. *Int J Pharm* 2014;**473**:64–72.
61. Amidon, G.E. Compaction and the properties of mixtures, presented at Compaction Simulation Forum, Cambridge, MA, November 13–14, 2012, http://people.clarkson.edu/~ccetinka/CompactionPresentation/1%20Amidon%20UMichigan.pdf.
62. Habib SY, et al. Is silicified wet-granulated microcrystalline cellulose better than original wet-granulated microcrystalline cellulose? *Pharm Dev Technol* 1999;**4**:431–7.
63. Guo M, Augsburger LL. Potential application of silicified microcrystalline cellulose in direct-fill formulations for capsule-filling machines. *Pharm Dev Technol* 2003;**8**:47–59.
64. Li JX, Wu XY, Zhou Y. Characterization of wet masses of pharmaceutical powders by triaxial compression test. *J Pharm Sci* 2000;**89**:178–90.
65. Rowe RC, Sheskey PJ, Cook WG, Fenton ME, editors. *Handbook of pharmaceutical excipients*. 7th ed. London: Pharmaceutical Press; 2012.
66. Rowe RC. Materials used in the film coating of oral dosage forms. In: Florence TA, editor. *Critical reports on applied chemistry*. Oxford: Blackwell Scientific; 1984. pp. 1–36.
67. Siepmann J, Peppas NA. Modeling of drug release from delivery systems based on hydroxypropyl methylcellulose (HPMC). *Adv Drug Deliv Rev* 2001;**48**(2-3):139–57.
68. Ozturk AG, et al. Mechanism of release from pellets coated with an ethyl cellulose-based film. *J Control Release* 1990;**14**:203–13.
69. Bodmeier R, Paeratakul O. The effect of curing on drug release and morphological properties of ethylcellulose pseudolatex-coated beads. *Drug Dev Ind Pharm* 1994;**20**(9):1517–33.
70. Dressman JB, et al. Circumvention of pH-dependent release from ethyl cellulose-coated pellets. *J Control Release* 1995;**36**(3):251–60.
71. Carlin B, Li J-X, Felton LA. In: McGinity, Felton, editors. *Pseudolatex dispersions for controlled drug delivery, aqueous polymeric coatings for pharmaceutical dosage forms*. 3rd ed. New York, NY: Informa Healthcare; 2008.
72. Celik M. Compaction of multiparticulate oral dosage forms. In: Ghebre- Sellassie I, editor. *Multiparticulate oral drug delivery*. New York, NY: Marcel Dekker; 1994. pp. 181–215.
73. Theeuwes F. Elementary osmotic pump. *J Pharm Sci* 1975;**64**:1987–91.
74. Santus G, Baker RW. Osmotic drug delivery: review of the patent literature. *J Control Release* 1995;**35**:1–21.
75. Van Savage G, Rhodes CT. The sustained release coating of solid dosage forms: a historical review. *Drug Dev Ind Pharm* 1995;**21**:93–118.
76. Thombre AG, Cardinal JR, DeNoto AR, Gibbes DC. Asymmetric membrane capsules for osmotic drug delivery II. In vitro and in vivo drug release performance. *J Control Release* 1999;**57**:65–73.
77. Hilton AK, Deasy PB. Use of hydroxypropyl methylcellulose acetate succinate in an enteric polymer matrix to design controlled-release tablets of amoxycillin trihydrate. *J Pharm Sci* 1993;**82**:737–43.
78. Takeichi Y, et al. Combinative improving effect of increased solubility and the use of absorption enhancers on the rectal absorption of uracil in beagle dogs. *Chem Pharm Bull* 1990;**38**:2547–51.
79. Xie T, Taylor LS. Dissolution performance of high drug loading celecoxib amorphous solid dispersions formulated with polymer combinations. *Pharm Res* 2015. First online: 12 November 2015.
80. Oppong F, et al. Microrheology and structure of a yield-stress polymer gel. *Phys Rev E Stat Nonlin Soft Matter Phys* 2006;**73**(4–1): 041405/1– 041405/9.
81. Khan GM, Jiabi Z. Formulation and in vitro evaluation of ibuprofen—Carbopol 974P-NF controlled release matrix tablets III: influence of coexcipients on release rate of the drug. *J Control Rel* 1998;**54**(2):185–90.
82. Efentakis M, et al. Dimensional changes, gel layer evolution and drug release studies in hydrophilic matrices loaded with drugs of different solubility. *Int J Pharm* 2007;**339**(1–2):66–75.
83. Llabot JM, et al. Drug release from carbomer:carbomer sodium salt matrices with potential use as mucoadhesive drug delivery system. *Int J Pharm* 2004;**276**:59–66.
84. Vaidya AP, et al. Protective effect of Carbopol on enzymatic degradation of a peptide-like substrate I: effect of various concentrations and grades of Carbopol and other reaction variables on trypsin activity. *Pharm Dev Technol* 2007;**12**(1):89–96.
85. Nollenberger K, Albers J. Poly(meth)acrylate-based coatings. *Int J Pharm* 2013;**457**:461–9.

86. Akhgari A, et al. Combination of time-dependent and pH-dependent polymethacrylates as a single coating formulation for colonic delivery of indomethacin pellets. *Int J Pharm* 2006;**320**:137–42.
87. Arno EA, et al. Eudragit NE30D based metformin/gliclazide extended release tablets: formulation, characterisation and in vitro release studies. *Chem Pharm Bull (Tokyo)* 2002;**50**:1495–8.
88. Zhu Y, et al. Influence of thermal processing on the properties of chlorpheniramine maleate tablets containing an acrylic polymer. *Pharm Dev Technol* 2002;**7**:481–9.
89. Pina MF, Zhao M, Pinto JF, Sousa JJ, Craig DQ. The influence of drug physical state on the dissolution enhancement of solid dispersions prepared via hot-melt extrusion: a case study using olanzapine. *J Pharm Sci* 2014;**103**:1214–23.
90. Elder DP. Pharmaceutical Applications of Ion-Exchange Resins. *J Chem Educ* 2005;**82**:575–87.
91. Guo X, Chang RK, Hussain MA. Ion-exchange resins as drug delivery carriers. *J Pharm Sci* 2009;**98**:3886–902.
92. Durig, T. et al. Importance of structure-function relationships and particle size distribution in optimizing powder flow and compactibility of directly compressible ethylcellulose. Ashland Pharmaceutical Technology Report, PTR-38; 2007.

This page intentionally left blank

CHAPTER

8

Interfacial Phenomena

E. Hadjittofis[1], *S.C. Das*[2], *G.G.Z. Zhang*[3] *and J.Y.Y. Heng*[1]

[1]Imperial College London, South Kensington Campus, London, United Kingdom
[2]University of Otago, Dunedin, New Zealand [3]Abbvie Inc., North Chicago, IL, United States

8.1 INTERFACES

In the pioneering work "Graphical methods in the thermodynamics of fluids" by Prof. J. Willard Gibbs,[1] he makes the argument that:

> It may occur, however, in the volume-entropy diagram that the same point represents two different states of the body. This occurs in the case of liquids which can be vaporized.

This work, the first integrated mathematical description of interfacial phenomena, sparked the development of a new research area dealing with the phenomena at the boundaries between phases.

The interface is the boundary between any two phases.[2] This definition implies that the two phases can be of the same state or even to be identical. For the sake of convenience, interfacial interactions are distinguished by the states of matter of two contacting phases, solid (S), liquid (L) and vapor (V). Therefore, six possible interfaces arise as a result of the combination of contacting phases: S-S, S-L, S-V, L-L, L-V, V-V. However, the existence of an interface between two vapor phases is unlikely, thus, it is omitted in this chapter. Additionally, the L-V interface is of less practical importance to solid dosage forms. Therefore, it will not be included in this chapter either. In all succeeding discussions, it is important to keep in mind the dynamic nature of these interfaces; for example, when a solid particle is dissolving in a liquid medium, the S-L interface decrease upon dissolution until it vanishes.

S-S interactions (commonly known as contact mechanics): Contact mechanics describe the cohesive and adhesive properties of particles implicating their processability during manufacturing. In particular, S-S interface influences the flow of powders,[3,4] mixing, and blending operations,[5,6] including coating and deaggregation processes.[7] Furthermore, milling and micronization processes,[8–12] which are very important for the size reduction of active pharmaceutical ingredients (APIs) for dry powder inhalers (DPIs), will have an impact on the API-excipient S-S interactions. Formulation processes such as fluidization and tableting are subjected to the effects of S-S interfaces as well.

S-L interactions: These are of particular importance in operations like wet milling and wet granulation,[13] where liquids are used to enhance the process efficiency. Interfacial properties will determine the wetting of solids in liquids, an important step in the formation of granule nuclei leading to successful granulation processes. Solid–liquid interactions will also determine the dissolution behavior of solid dosage forms, in vitro and in vivo.[14]

S-V interactions: Moisture uptake can lead to the accumulation of liquid on solids either in the form of bound or unbound water. The adsorption mechanisms are susceptible to changes in solid properties such as surface energy, porosity and physical form (amorphous vs crystalline), as well as the changes in the environment. Morphological changes and recrystallization[15,16] have been reported as a result of moisture uptake. Moisture removal is just as important as moisture uptake. It is widely reported that the formation of polymorphs through dehydration/desolvation hinges on drying conditions and dehydration/desolvation mechanisms.[17,18]

L-L interactions: Emulsification processes are employed in delivering insoluble drugs and in preparing drug delivery carrier particles. These are two important examples of the application of L-L interactions.

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00008-X

8.2 FUNDAMENTAL INTERMOLECULAR FORCES

In his Nobel Lecture on December 12, 1910,[19] Johaness Diderik van der Waals challenged Boyle's law, stating:

> As you are aware the two factors which I specified as reasons why a nondilute aggregate of moving particles fails to comply with Boyle's law are firstly the attraction between the particles, secondly their proper volume.

This statement has been interpreted mathematically by the well-known van der Waals equation of state:

$$\left(p + \frac{a_{\mathrm{VdW}}}{v^2}\right)(v - b_{\mathrm{VdW}}) = RT \tag{8.1}$$

where a_{VdW} describes the adhesive (van der Waals) forces between the molecules, b_{VdW} corresponds to the exclusion volume of a molecule of the fluid and the rest of the parameters have the same meaning as in the ideal gas law. As a_{VdW} and b_{VdW} tend to zero, the van der Waals equation reduces to the ideal gas law. This condition is in line with the fundamental assumptions of ideal gas law, assuming that the gas molecules are volumeless and molecules do not interact.

8.2.1 Van der waals forces

Van der Waals forces can be categorized, on the ground of the molecules involved in the interaction, to: Keesom forces, Debye forces, and London (dispersion) forces, summarized schematically in Fig. 8.1. For Keesom forces, two polarized molecules interact because of the inherent difference in charge distribution. In the case of Debye forces, a molecule with a permanent dipole induces charge redistribution to neighboring molecules with no dipole moments. Finally, London forces arise in molecules without permanent dipoles. The fluctuations on the electron cloud lead to temporary changes in the charge distribution inducing a charge redistribution to neighboring molecules. The mathematical formulation of all three components has the general form:

$$U(r) = -\frac{C}{r^6} \tag{8.2}$$

where C is a constant changing slightly for each component and r is the intermolecular distance.

Apart from van der Waals forces, there exists an interaction between electron poor and electron rich atoms, via the sharing of a lone pair of free electrons from the latter to the former, called *dipole-dipole interactions*. When a hydrogen atom is involved in this interaction, the interaction is called hydrogen bond. Due to its nature, it is characterized by directionality and short-range action, meaning that the electron poor and the electron rich sites should "face" one another at a close distance. Directionality leads to the formation of weak structures of molecules held together via the bonds formed through the sharing of lone pairs of free electrons. The effect of these bonds is affected by the proximity of the molecules. Thus, it changes depending on the state of the matter. In particular, as matter moves from vapor to liquid or solid states, the importance of them increases. The physical chemical property of ice is the most striking manifestation of the effects of these forces.

Additionally, Coulombic forces develop between charged atoms or ions. The nature (attractive or repulsive) and magnitude of the interaction depend on the sign (positive or negative), the ionic charge of the contributing elements, the distance, and the medium separating them. For instance, two oppositely charged ions exhibit attractive interactions, the magnitude of which is given by the following relation:

$$U(r) = \frac{Q_1 Q_2}{4\pi\varepsilon_0 r^2} \tag{8.3}$$

where Q_1 and Q_2 are the charges, r is the distance between them, ε_0 is the dielectric constant of the medium separating these ions.

8.2.2 Thermodynamics of dispersion forces

From the aforementioned van der Waals forces, London dispersion forces account for the majority of the van der Waals interactions. In the following section, two main approaches for the derivation of them will be discussed: Hamaker's and Lifshitz's approaches.

8.2.2.1 Hamaker's approach[20,21]

If a single spherical molecule is suspended above a flat, solid surface (Fig. 8.2), it can be assumed that the total interaction to be given by the summation of all the intermolecular interactions between the molecule and the flat surface. Assuming simple additivity

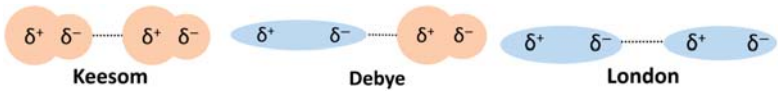


FIGURE 8.1 Schematic representation of the three components of van der Waals forces.

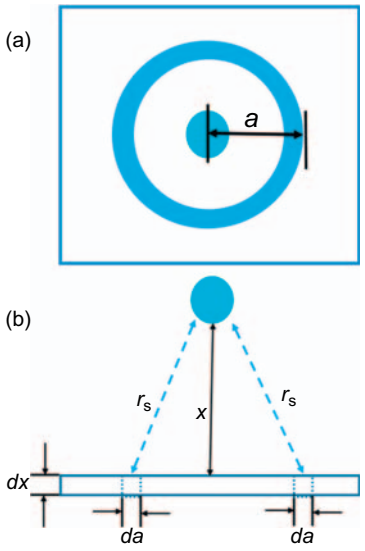


FIGURE 8.2 Schematic representation for the derivation of potential energy equation by Hamaker's approach (a) top view, (b) side view.

between them and n number of molecules on the surface then the potential energy, U, can be written as:

$$\frac{dU}{dn} = -\frac{3\alpha^2 h\upsilon}{4(4\pi\varepsilon_0)^2 r_s^6} \tag{8.4}$$

where r_s is the geometric term shown in Fig. 8.2, α is the polarizability, h stands for the Planck constant, υ is the frequency of fluctuation and ε_0 is the dielectric permittivity in vacuum.

The number of molecules, assuming a spherical shape, can be calculated further by the following equation:

$$dn = 2\pi\ \rho da\ dx \tag{8.5}$$

where ρ is the concentration of molecules on the surface.

Upon combining the two equations and integrating, the following relation for the potential energy can be obtained:

$$U = \int_H^\infty \int_0^\infty -\frac{3a^2 h\upsilon}{4(4\pi\varepsilon_0)^2 (x^2+\alpha^2)^3} 2\pi\ \rho d\alpha\ dx = \frac{3a^2 h\upsilon}{4(4\pi\varepsilon_0)^2}\frac{\pi\rho}{6H^3} \tag{8.6}$$

As can be seen in Fig. 8.2, α is the radius of the flat solid surface, x is the variable used to describe the distance between the molecule and the surface and finally H is the actual distance between the molecule and the surface.

For a system of two parallel plates the number of molecules is given simply by $dn = \rho \cdot dh$, so the potential is given by:

$$U = \frac{3a^2 h\upsilon}{4(4\pi\varepsilon_0)^2}\frac{\pi\rho^2}{12H^2} \tag{8.7}$$

Hamaker conducted pairwise summation of the dispersion energies and proposed that the interaction between two similar macroscopic bodies at short distances can be given by:

$$A_{ii} = \beta\pi\rho^2 \tag{8.8}$$

where β stands for the pairwise interaction coefficient.

where A_{ii} is the Hamaker constant.

As such, the dispersive intermolecular energy or the free energy of interaction between two parallel plates can be rewritten as

$$U = \Delta G = -\frac{A_{ii}}{12\pi r_s^2} \tag{8.9}$$

where r_s is the interplate separation.

In the case of two bodies of different materials, Berthelot's principle states that the interaction can be estimated by the geometric mean of the interaction between each of the same materials. Thus, the Hamaker constant for their interactions is written as:

$$A_{ij} = \sqrt{A_{ii}A_{jj}} \tag{8.10}$$

where A_{ii} and A_{jj} are the Hamaker constants of each material in vacuum.

Similarly for two bodies (1, 2) interacting in a third medium (3), the Hamaker combination rule gives the following expression:

$$\begin{aligned} A_{132} &\approx A_{12} + A_{13} - A_{13} - A_{23} \\ &= \left(\sqrt{A_{11}} - \sqrt{A_{33}}\right)\left(\sqrt{A_{22}} - \sqrt{A_{33}}\right) \end{aligned} \tag{8.11}$$

8.2.2.2 *Lifshitz's approach*[22]

Lifshitz avoided the problems associated with additivity proposed by the earlier approaches of Hamaker. Lifshitz's approach ignores the atomic structure and treats the forces between large bodies as continuous media. The Lifshitz theory is derived from dielectric constants and refractive indices. Nonetheless, the full derivation is out of the scope of this chapter as it requires extensive understanding of quantum mechanics.[23,24] The Hamaker constants, for unlike bodies interacting in a medium, evaluated by the Lifshitz approach are given by

$$\begin{aligned} A_{132} = {} & \frac{3}{4}kT\left(\frac{\varepsilon_1-\varepsilon_3}{\varepsilon_1+\varepsilon_3}\right)\left(\frac{\varepsilon_2-\varepsilon_3}{\varepsilon_2+\varepsilon_3}\right) \\ & + \frac{3h\upsilon}{8\sqrt{2}}\frac{\left(n_1^2-n_3^2\right)\left(n_2^2-n_3^2\right)}{\sqrt{\left(n_1^2+n_3^2\right)\left(n_2^2+n_3^2\right)}\left(\sqrt{\left(n_1^2+n_3^2\right)}+\sqrt{\left(n_2^2+n_3^2\right)}\right)} \end{aligned} \tag{8.12}$$

where n_i and ε_i stand for the refractive index and dielectric constant of material i, respectively.

The main limitation to the Lifshitz approach is that the theory is based on the continuum theory, which is true only when the interacting surfaces are farther apart than their molecular dimensions. The practical use of this theory is dependent on the availability of the dielectric constants for the materials of interest.

8.3 THERMODYNAMICS OF PARTICLES IN ELECTROLYTE SOLUTIONS

8.3.1 DLVO theory[25,26]

Let's now consider an electrolyte containing solution with negatively charged particles in it, with D being the distance between two particles. For the sake of convenience, it is first assumed that only counterions (positive ions with respect to the negatively charged surfaces) exist in the solution. The surfaces of the particles are treated as flat since in a microscopic point of view curvature effects can be neglected. The chemical potential of the ions can be written as:

$$\mu = ze\psi + kT\log(\rho) \tag{8.13}$$

where z corresponds to a particular valency number, ψ stands for the electrostatic (zeta) potential, ρ is the number density of ions with valency z at a distance x from the midpoint (Fig. 8.3).

The repulsive forces between charged particles in a solution can be described according to the following equation:

$$W_{\text{rep}}(D) = 2\pi Rk\varepsilon_0\varepsilon_{\text{R}}\psi_0^2\exp(-\kappa D) \tag{8.14}$$

where R is the radius of the sphere, ε_0 is the electric permittivity in vacuum, ε_{R} is the dielectric constant of the liquid, κ is a constant characterizing the width of the electrical double layer, and D is the distance between two particles.

8.3.2 Steric stabilization of particles

Polymer adsorption on particles surface is used to stabilize colloidal systems. The adsorbed polymers create a layer around the particle. When two particles come close, their polymer layers interact and the arising steric forces prevent coagulation, stabilizing the system. Both enthalpic and entropic phenomena contribute to the steric potential, the former associated with the elastic deformation of the polymer layer and the latter with the osmotic phenomena in the polymer layer. The enthalpic and entropic contributions act additively, as shown in the following equation (Fig. 8.4):

$$W_{\text{steric}} = W_{\text{enthalpic}} + W_{\text{entropic}} \tag{8.15}$$

The interaction between the polymer chains increases the local concentration of polymer inducing

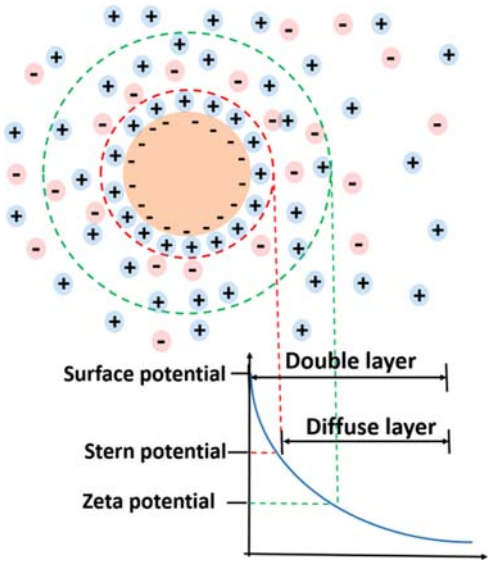


FIGURE 8.3 Schematic showing the double layer around a negatively charged particle and the corresponding graph showing the change in the magnitude of the potential at different distances away from the surface of the particle.

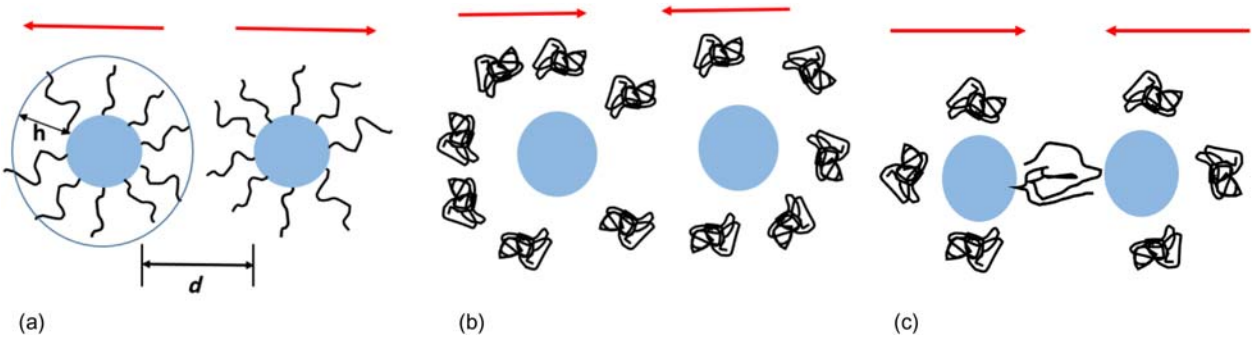


FIGURE 8.4 Illustration of three possible interactions between polymers and particles (the red arrows show whether the particles come closer or move further away): (a) Particles coated with sufficient amount of polymer repel each other, stabilizing the system, (b) If the polymer does not adsorb on the particles, the steric interactions are not sufficient, then there are not enough steric interactions and particles attract each other, and (c) If the amount of adsorbed polymer is not sufficient, bridges are formed attracting the particles.

an osmotic force driving solvent molecules to diffuse toward the gap between the two particles, forcing the two particles to disjoin. Moreover, the interaction between the polymer chains reduces the conformational entropy of the polymer, increasing the free energy of the system enhancing its stability. A coherent mathematical interpretation of this phenomenon can be given in terms of Fischer's work[27]:

$$W_{\text{steric}} = 2\pi RkT\Gamma^2\left(\mathrm{N_A}\left(\frac{V^2_{\text{polymer}}}{V_{\text{solvent}}}\right)\left(\frac{1}{2}-\chi\right)S_{\text{osmotic}} + \frac{1}{\Gamma}S_{\text{entropic}}\right) \tag{8.16}$$

In this equation, Γ is the amount of adsorbed polymer, V_{polymer} and V_{solvent} are the molar volumes of the polymer and the solvent, respectively, and R is the radius of the particle. The term χ is the Flory–Huggins parameter discussed in Section 2.7, S_{osmotic} and S_{entropic} are geometric terms determined by the particle size and the polymer behavior of the polymer layer.

As can be seen, the strength of the solvent is driving the mixing component of the steric potential. If the solvent is strong, meaning that it can dissolve a large amount of material, then $\chi < 0.5$, the overall steric potential has constantly a positive value, favoring good dispersion of the particles. On the other hand, if $\chi > 0.5$, then there would be a competition between osmotic and entropic phenomena. Furthermore, molar-density appears to be quite important. In good solvent conditions, low density polymers enhance mixing, and the reverse happens for bad solvent conditions.

In the context set here, the total potential for sterically stabilized colloid particles is given by the summation of the entropic and enthalpic phenomena with the effects of van der Waals forces described earlier in this chapter:

$$\begin{aligned} W_{\text{total}} &= W_{\text{enthalpic}} + W_{\text{entropic}} + W_{\text{Van der Waals}} \\ &= 2\pi RkT\Gamma^2\left(\left(\frac{V^2_{\text{polymer}}}{V_{\text{solvent}}}\right)\left(\frac{1}{2}-\chi\right)S_{\text{osmotic}} + \frac{1}{\Gamma}S_{\text{entropic}}\right) \\ &\quad + W_{\text{Van der Waals}} \end{aligned} \tag{8.17}$$

If the thickness of the layer is not sufficient, then S_{osmotic} decreases rapidly as the interparticle distance increases. This can result in negative values of W corresponding to a flocculating system.

8.4 SURFACE TENSION AND SURFACE ENERGY

8.4.1 Fundamentals

In the bulk of a material, at any state, molecules interact symmetrically with the rest as shown in Fig. 8.5.

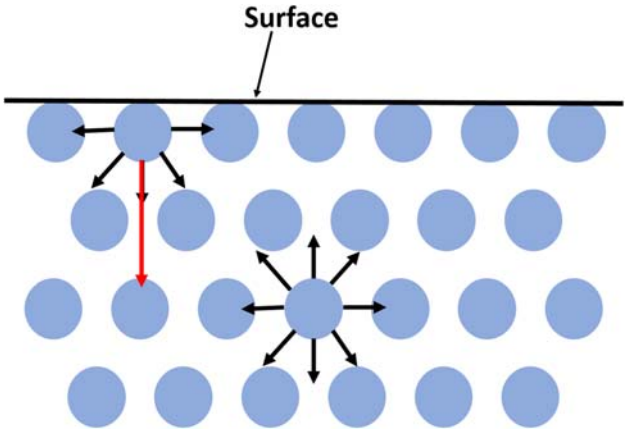


FIGURE 8.5 Interactions for molecules in bulk and molecules on the surface of a material. For the resultant forces shown on one of the surface molecules with a red arrow, there is an equal magnitude and opposite direction force corresponding to the surface tension.

However, the molecules on the surface of the bulk exhibit anisotropic interactions leading to a net force toward the bulk of the material. This force imbalance gives rise to the concept of surface tension (for liquids) or surface energy (for solids). This force imbalance is the energy needed to create a unit surface area of a material.

Consider the interface between a liquid and a vapor having a height h and surface area A; then its volume is given according to the relation $V = h \cdot A$. The work needed (W) to create a unit area of this interface should be the sum of the effects of surface energy and volume expansion:

$$dW = \gamma\, dA - P\, dV \tag{8.18}$$

where γ is the surface tension of the liquid and P is the pressure of the system.

The internal energy of the system is then given by:

$$dU = dQ + dW = TdS + \gamma\, dA - P\, dV \tag{8.19}$$

Furthermore, the expression for Gibbs free energy comes as follows:

$$\begin{aligned} dG &= dH - d(TS) = dU + d(PV) - d(TS) \\ &= TdS + \gamma\, dA - P\, dV + PdV + VdP - SdT - TdS \\ &= \gamma\, dA + VdP - SdT \end{aligned} \tag{8.20}$$

If the system is at constant temperature and pressure, then the Gibbs free energy equation collapses further, to

$$dG = \gamma\, dA \Longrightarrow \gamma = \left(\frac{\partial G}{\partial A}\right)_{\text{T,P,V}} \tag{8.21}$$

The derived relation proposes a "thermodynamic" definition for surface tension to be the change in Gibbs free energy per change in surface area at constant pressure, temperature and volume.

If constant pressure and surface area are assumed, then the change in Gibbs free energy is:

$$dG = -SdT \tag{8.22}$$

Since it is known that Gibbs free energy and surface tension are identical, then this can be rewritten as

$$S = -\left(\frac{dG^s}{dT}\right)_P = -\left(\frac{d\gamma}{dT}\right)_P \tag{8.23}$$

The negative sign in the relation indicates that the free energy at the surface results in attractive forces between materials. For two separate and different bodies, the work associated with holding them together at an interface is termed the adhesive work, W_A while the work associated to holding together bodies of the same kind is called the cohesive work, W_c.

8.4.2 Surface energy components

As previously mentioned, the surface energy of a material is the energy required to create a unit area of the material. In this sense, it can be described in physicochemical terms, according to Fowkes' hypothesis,[28] as the sum of intermolecular forces the system is composed of. Mathematically this corresponds to Eq. 8.24, where n stands for the major surface energy components (dispersion forces, hydrogen bonds, dipole-dipole interactions, ion-dipole interactions, etc.). This equation is usually written in a simplified form as described by Owens–Wendt,[29] with one term accounting for the apolar (Lifshitz–van der Waals) interactions, γ^{LW}, and one term accounting for the polar interactions, γ^P:

$$\gamma = \sum_n \gamma^n = \gamma^{LW} + \gamma^P \tag{8.24}$$

The polar component comprises acid-base interactions and hydrogen bonds. On the other hand, the Lifshitz–van der Waals components not only include dispersion (London) forces,[30] but also the forces induced by the orientation of the interacting molecules (Keesom forces) and by the interactions between a polar and a nonpolar molecule (Debye interaction).[21] Nevertheless, for convenience, γ^{LW} is called the *dispersive component of surface energy*.

The manifestation of surface energy in the macroscopic world can be seen, as mentioned, through the work of adhesion. Let's assume two dissimilar surfaces, i and j, trying to join each other. The total work (W_{ij}) needs to be done is given by

$$W_{ij} = \gamma_i + \gamma_j - \gamma_{ij} \tag{8.25}$$

In this equation, γ_i and γ_j, correspond to the surface energy of the surfaces i and j, respectively, where γ_{ij} stands for the surface energy of the interface. To calculate the value of γ_{ij}, the fundamental equation of surface energy is applied in the context of an interface.

$$\gamma_{ij} = 2\left(\sqrt{\gamma_i^{LW}\gamma_j^{LW}} + \sqrt{\gamma_i^P\gamma_j^P}\right) \tag{8.26}$$

For this document, a geometric mean approach (the Girifalco-Good equation) is selected to calculate the value of the surface energy of the interface.[31] However, another model has been suggested by Wu, using a harmonic mean approximation in the calculations as shown in the following equation[32]:

$$\gamma_{ij} = 4\left(\frac{\gamma_i^{LW}\gamma_j^{LW}}{\gamma_i^{LW}+\gamma_j^{LW}} + \frac{\gamma_i^P\gamma_j^P}{\gamma_i^P+\gamma_j^P}\right) \tag{8.27}$$

The polar component of Eq. 8.27 can be further analyzed according to van Oss-Chaudury-Good (vOCG) method.[24] The polar components are separated to take in account the electron acceptor (γ^+) and donor (γ^-) interactions between the two surfaces, giving us the following:

$$\gamma_{ij} = 2\left(\sqrt{\gamma_i^{LW}\gamma_j^{LW}} + \sqrt{\gamma_i^+\gamma_j^-} + \sqrt{\gamma_i^-\gamma_j^+}\right) \tag{8.28}$$

In cases where polar interactions are negligible, the polar components get a zero value. When the surfaces i and j are the same, then $\gamma_i = \gamma_j$. We refer to this situation as energy of cohesion.

8.4.2.1 *Acid-base interactions*

As has been discussed in the previous section, the vOCG approach accounts for the short-range forces in terms of acid (γ^+) and basic (γ^-) parameters. The hydrogen bonding geometric mean approximation for compounds like ethers, esters and aromatics was challenged.[33] This is due to the hydrogen bonding component of the work of adhesion of these compounds with hydrogen donors, even though their structure does not permit the formation of hydrogen bonds (corresponding to $\gamma_{AB}^{hydrogen} = 0$). Instead of this erroneous approximation, Fowkes and Mostafa proposed the following relation for the acid-base component of the work of adhesion as follows:

$$W^{AB} = -f * N^{AB} * \Delta H^{AB} \tag{8.29}$$

In this relation, ΔH^{AB}stands for the enthalpy of adhesion, N^{AB} for the number of acid-base pairs per unit area and f for a conversion constant used to normalize the units. This equation implies experimental calculation of ΔH^{AB}. In this direction, two different models have been proposed.

The first, developed by Drago,[34] is a combination of Mulliken's work on molecular complexes and the Hard-Soft Acid-Base theory.[35] According to this model, enthalpy can be measured in terms of the tendency of

the acidic or the basic components of electrostatic interactions and covalent bonding. Mathematical formulation of it is given in the following form:

$$\Delta H^{AB} = -(E_A E_B + C_A C_B) \tag{8.30}$$

To avoid any misconception, it is important to specify that for this equation the two terms with subscript A correspond to the electrostatic (E) and covalent (C) parameters of the acidic components, respectively, and those with B subscript to the same terms of the alkaline component. The ratio of the two terms, E and C, of each component defines the hardness of the corresponding component. Iodine is used as the reference, giving a ratio equal to 1. The validity of this method for solids with a very high or very small hardness ratio has been disputed.

The second method for the determination of the acid-base component of the surface energy, was proposed by Gutmann.[36] This theory is represented mathematically as follows:

$$\Delta H^{AB} = -(K_A DN + K_B AN^*) \tag{8.31}$$

In this equation, the K's are parameters characterizing the acidic (A) and alkaline (B) capacity of the solid, respectively. In addition, DN and AN^* stand for the donor and the corrected acceptor number, respectively. The first one, DN, is a parameter measuring the electron donor characteristics. It is measured based on the heat of mixing of the compound in a solution of antimony pentachloride and dichloroethane. The corrected acceptor number (AN^*) is a parameter calculated indirectly from the acceptor number (AN). Acceptor number is measured based on the shifts the compound causes on the P-NMR spectrum of tri-ethyl phosphine. The correction was introduced to account for the effects of van der Waal forces on the shift. Using solvent probes with different acid-base properties, a plot of $\Delta H^{AB}/AN^*$ against DN/AN^* can be constructed. The slope would correspond to K_A and the interception with the y-axis is K_B.

8.4.3 Fundamentals of self-assembly of soft Structures

Poorly water-soluble drug substances, as well as those degrading in water, are often formulated with the aid of soft structures[37] such as microemulsions.[37,38] These structures are primarily constructed by amphiphilic molecules, the structure of which encompasses hydrophobic and hydrophilic parts, such as surfactants and copolymer. The main aim of this section is to provide an overview of the thermodynamic principles underlying self-assembly phenomena creating soft structures out of amphiphilic molecules. The abundance of possible structures of soft structures arises from the anisotropic nature of the amphiphiles and from the spatial arrangement of the functional groups contributing to self-assembly via attractive interactions.

Considering a system of amphiphiles creating soft structures comprising N monomers, the chemical potential of the system can be written as follows:

$$\begin{aligned}\mu &= \mu_1^o + kT\log(X_1) = \mu_2^o + \frac{1}{2}kT\log\left(\frac{1}{2}X_2\right)\\ &= \ldots = \mu_N^o + \frac{1}{N}kT\log\left(\frac{1}{N}X_N\right)\end{aligned} \tag{8.32}$$

where μ_i is the chemical potential of species i, X_i stands for the concentration of i.

Self-assembly is taking place in solutions, in case of equilibrium in this solution the chemical potential of the micelles should be equal to the chemical potential of the free amphiphiles in solution. From the thermodynamic derivation, it is clear that for stable micellization the following condition should be fulfilled:

$$\Delta\mu = \mu_{\text{micelle, N}} - \mu_{\text{solution}} < 0 \tag{8.33}$$

So, as the size of the structure increases, the total chemical potential of the system decreases. However, it is expected that the spatial conformation of the soft structures affect the chemical potential of the system. In this direction, it is important to be able to distinguish the implications associated with geometry, from first principles. Starting from simple one-dimensional (1D) structures, such as rods, comprising of N components it is implied that the change in chemical potential decreases linearly with increasing value of N. Similarly, for two-dimensional structures like circles and sheets, the decrease is proportional to $N^{1/2}$, and for three-dimensional micelles to $N^{1/3}$.

$$\Delta\mu_N = \frac{\alpha kT}{N}(1\ D) \tag{8.34}$$

$$\Delta\mu_N = \frac{\alpha kT}{N^{\frac{1}{2}}}(2\ D) \tag{8.35}$$

$$\Delta\mu_N = \frac{\alpha kT}{N^{\frac{1}{3}}}(3\ D) \tag{8.36}$$

Term α in these equations is a positive constant related to the different geometries.

The changes in system properties occurring above critical micelle concentration (CMC) are used to study micelle formation. At concentrations below CMC, there is a linear increase in monomer concentration as total concentration increases. Beyond CMC, however, the monomer concentration does not change any more. All added molecules self-assemble in the form of micelles, leading to the changes in a range of properties of the system. This sharp change is reflected in changes in a wide range of physical properties, including but not limited to, osmotic pressure, turbidity and surface

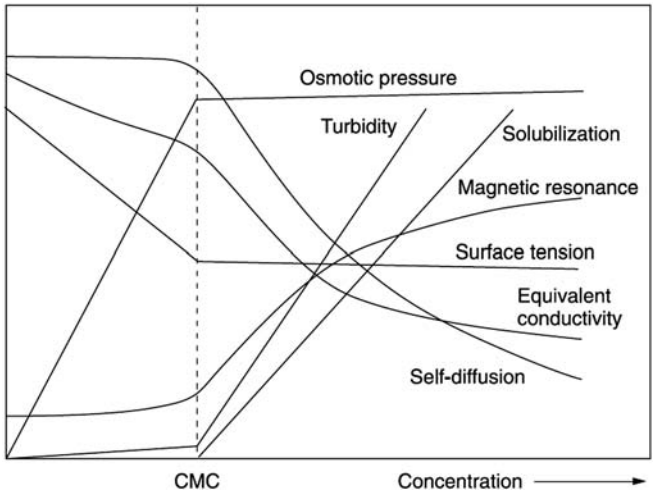


FIGURE 8.6 Change in physicochemical properties of system as a function of surfactant concentration.

tension, as shown in Fig. 8.6. Thus, the micelle concentration can be assessed indirectly by measuring these properties with well-established appropriate techniques.

Nonetheless, properties such as turbidity are affected by the size distribution of the aggregates, as real systems do not exhibit monodispersed aggregates, with a certain value of N, but structures with a range of sizes. To appreciate the relative abundance of different aggregate sizes, it is important to understand the thermodynamic principles governing aggregation of structures with different size. For convenience, in terms of mathematical manipulations, 1D structures comprising an N number of monomers is assumed.

Starting from the fundamental thermodynamic equations for self-assembly leading to

$$\mu_1^{\circ} + kT\log(X_1) = \mu_N^{\circ} + \frac{1}{N}kT\log\left(\frac{1}{N}X_N\right) \rightarrow X_N = N\left(X_1\exp\left(\frac{\mu_1^{\circ}-\mu_N^{\circ}}{kT}\right)\right)^N \quad (8.37)$$

Appropriate mathematic manipulation of these equations, a parabolic equation can be derived describing the change in chemical potential of a system in terms of assemblies comprising N number of micelles and the average number of micelles in each assembly (M):

$$\mu_N^{\circ} - \mu_M^{\circ} = \Lambda(\mathrm{N}-\mathrm{M})^2 \quad (8.38)$$

where Λ is a geometric constant.

8.5 THERMODYNAMICS OF WETTING

8.5.1 Fundamentals

As shown in Section 8.4, the surface tension of a liquid can be defined as the change in Gibbs free energy of a system of changing surface area at constant temperature, pressure and volume. Consider a liquid i spreading over a solid surface j, new interfaces ij are created during this wetting process. This process is interpreted mathematically as follows:

$$\partial G = \left(\frac{\partial G}{\partial A_i}\right)_{T,P} dA_i + \left(\frac{\partial G}{\partial A_j}\right)_{T,P} dA_j + \left(\frac{\partial G}{\partial A_{ij}}\right)_{T,P} dA_{ij} \quad (8.39)$$

Using the definition of surface energy/tension, this equation can be written as

$$-\left(\frac{\partial G}{\partial A_i}\right)_{T,P} = S_{ij} = \gamma_j - \gamma_i - \gamma_{ij} \quad (8.40)$$

where S_{ij} is the spreading coefficient, also related to the work of cohesion and adhesion. Since the work of cohesion can be expressed in terms of surface tension energy by the equation $W_C = 2\gamma_i$, then the work of adhesion can be calculated as follows:

$$S_{ij} = W_A - W_C = \gamma_j - \gamma_i - \gamma_{ij} \quad (8.41)$$

$$W_A - 2\gamma_i = \gamma_j - \gamma_i - \gamma_{ij} \quad (8.42)$$

$$W_A = \gamma_j + \gamma_i - \gamma_{ij} \quad (8.43)$$

8.5.2 Experimental techniques

8.5.2.1 Sessile drop contact angle

Sessile drop contact angle measurements employ the fundamentals of liquid spreading to determine the surface energy of a solid. It is grounded on the observations presented in Young's pioneering work on the cohesion of fluids[39] and can be summarized in the following statement, adapted from that work:

> It is necessary to premise one observation, which appears to be new, and which is equally consistent with theory and with experiment; that is, that for each combination of a solid and a fluid, there is an appropriate angle of contact between the surfaces of the fluid exposed to the air, and to the solid.

The fundamentals of it are described with the aid of Fig. 8.7. Consider a droplet resting on a solid surface. If the mass of the droplet is not big, the gravity effects can be omitted and the forces acting on the system are as shown, γ_{SL} stands for the solid–liquid interfacial energy, γ_{SV} is the interfacial energy between the solid and the air (the one that needs to be determined usually), γ_{LV} is the interfacial energy between the liquid and the air (commonly known as surface tension) and θ_c is the contact angle. The effects of the three-phase interfacial energy have also been neglected. Assuming that the droplet is in equilibrium, the resultant force on the system would be zero. So, Young's equation[40] can be written as

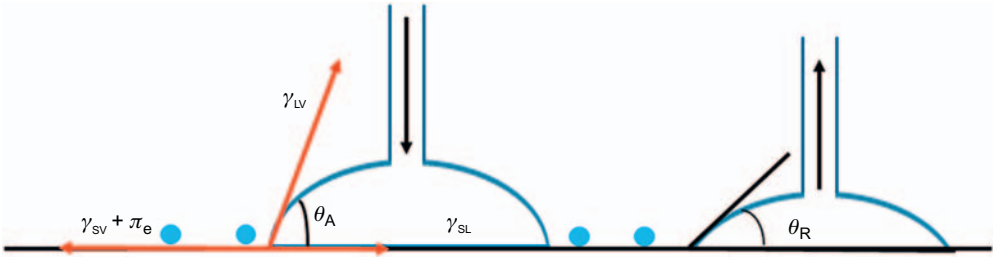


FIGURE 8.7 Schematic representation of the concept of advancing and receding contact angle measurements, indicating the relevant surface tensions, according to Young's equation and the spreading pressure.

$$\gamma_{SL} = \gamma_{SV} - \gamma_{LV}\cos\theta_c \tag{8.44}$$

However, experience leads to the conclusion that droplets do not have a unique contact angle, corresponding to a unique equilibrium position but they experience a wide range of contact angles. Studies conducted on this topic reveal that this broad spectrum is related to the structural and chemical heterogeneity of the material.[21,41,42] The effects of structural heterogeneity on wetting have been investigated in depth, with an example application in the development of biomimetic surfaces.[43–46] Theoretically, a clean and perfectly smooth ideal surface would lead to a single equilibrium position. Contact angles on nonideal surfaces can exhibit a maxima or minima, referred to as the *advancing* and *receding contact angle,* respectively, shown in Fig. 8.7.[47]

Experimentally, the droplet is placed on the surface and is inflated through pumping liquid into it. Simultaneously, measurements of the advancing contact angle are taken at regular, constant time intervals. It is important not to overinflate the droplet, because the effects of gravity become significant and it may collapse (the stick-slip phenomenon). When the droplet reaches a sufficiently large volume, liquid is removed from the droplet and simultaneously receding contact angle is measured. The average of the advancing (θ_A) and receding (θ_R) contact angles is often calculated from Eqs. 8.45–8.47 to calculate the "real" contact angle (θ_c). The difference between the advancing and receding contact angle is known as the contact angle hysteresis.

$$\theta_c = \arccos\left(\frac{r_A\cos\theta_A + r_R\cos\theta_R}{r_A + r_R}\right) \tag{8.45}$$

where r_A and r_R are given by:

$$r_A = \left(\frac{\sin^3\theta_A}{2 - 3\cos\theta_A + \cos^3\theta_A}\right)^{\frac{1}{3}}; \tag{8.46}$$

$$r_R = \left(\frac{\sin^3\theta_R}{2 - 3\cos\theta_R + \cos^3\theta_R}\right)^{\frac{1}{3}}; \tag{8.47}$$

Semiempirical models based on equations of state have been proposed for the measurement of the contact angle. These models introduce a new term, based on the degrees of freedom of the system, to account for the deviations from Young's equation. Li and Neumann[48,49] have proposed the following equation:

$$\cos\theta_c = -1 + 2\sqrt{\frac{\gamma_{SV}}{\gamma_{LV}}}\exp\left(-\beta\left(\gamma_{LV} - \gamma_{SV}\right)^2\right) \tag{8.48}$$

The term β has been determined experimentally, by fitting of a large number of experiments. This equation has been successfully used in polymeric systems. Nevertheless, its applicability on pharmaceutical solids has been disputed.[50]

8.5.2.2 *Beyond the sessile drop measurements*

Sixty years after Young's work, Wilhelmy[51] developed an experimental approach for the determination of the contact angle using the immersion of a solid into a liquid and measuring the force exerted on the solid due to capillary forces.

Because of the experimental setup, this technique is particularly useful for the measurement of the surface energy of thin films or fibers, but not very suitable for particulate materials. However, studies exist trying to apply this technique on compressed pharmaceutical powder disks and on glass slides with adhering pharmaceutical powders.[52] Glass slides were first sprayed with a liquid adhesive, and powder was adhered on the slide by shaking slides briefly in a container of the test powder. Theophylline and caffeine powders were studied in this work. Dove and coworkers found that the surface energy values obtained from the powder plates were more realistic compared to results obtained from compressed powder disks. This difference was attributed to possible changes in the surface properties of the powders during the preparation of the compacted powder disks. The effect of compression on surface chemistry has always been a point of concern for samples prepared in this manner. While the adhesion of powders onto slides overcomes some difficulties faced by compacted samples, it introduces other variables such as surface roughness as well as a potential contribution by the adhesive to the results obtained.

Washburn, Bartell, and Osterhof[53] use the same notion as Wilhelmy to develop a method measuring the surface energy of particulate materials using the rate of penetration of a liquid in the powder bed. Nevertheless, this approach includes numerous assumptions based on idealized conditions and may not be applicable to real particulate materials, especially for porous material.[54–56]

8.5.2.3 *Effects of surface roughness*

Advancing and receding contact angles are a result of surface chemical and structural heterogeneity, and have been first addressed by Cassie[57,58] and Wenzel,[59] respectively. Wenzel suggested that the measured contact angle can be correlated with the actual one based on a roughness parameter ($r_w > 1$).

$$r_w \cos(\theta_{actual}) = \cos(\theta_{measured}) \tag{8.49}$$

As can be seen from the equation, if the surface energy of the material corresponds to a $\theta > 90^\circ$ then increasing the roughness parameter (rougher surface) would result in an increase in the measured contact angle. The opposite happens with surfaces where $\theta < 90^\circ$.

Further studies lead to the development of the Cassie–Baxter equation, describing the contact angle on composite faces:

$$\cos(\theta) = \sum_{i=1}^{n} f_i \cos(\theta_i) \tag{8.50}$$

where f_i stands for the relative surface coverage of component i on the surface of interest and θ_i is the contact angle of the probe on pure component i. There is suggestion that advancing-receding contact angle measurements can be indicative of the surface energy, even for the presence of concentrated impurities. However, it seems that these kinds of measurements can be hardly extended to particulate materials owing to experimental limitations related with the wetting perimeter. Furthermore, the anisotropic nature of particulate materials, imposing energetic anisotropy and heterogeneous distribution of roughness, limits the accuracy of this type of measurements. Vapor sorption measurements,[60] using the notions developed by Anvir and Pfeifer on the fractal nature of surface roughness,[61,62] have been used to assess the roughness of powder materials, including pharmaceutical solids.[60] These sort of measurements can potentially be helpful for the determination of surface energy in terms of Wenzel equation.

Ambient humidity condensates on asperities forming capillaries, on the ground of Kelvin equation. These capillaries influence the cohesion and adhesion behavior of particles. Surface force apparatus and atomic force microscopy have been used extensively to characterize the effect of asperities in interparticle interactions.[63–66] Nonetheless, these measurements cannot, usually, account for a statistically significant number of particles and the corresponding interactions. Adsorption based techniques, such as finite dilution inverse gas chromatography, have been used[67,68] to overcome these problems. Ghoroi has investigated extensively the properties of coated and uncoated pharmaceutical particles. Some of the key findings are that increasing roughness leads to reduced moisture uptake leading to higher bulk density and less change of volume upon compression. Most importantly, it seems that coated pharmaceutical particles exhibit less sensitivity to their properties upon changing in relative humidity (RH).[69–71]

8.5.3 Implications of solid–liquid interfaces

8.5.3.1 *Interfacial thermodynamics in dissolution*

Dissolution of particles, surrounded by a solvent, is a multistep process during which a homogeneous solute–solvent/solution mixture is created with the mass transfer from the particles to the bulk. Thus, both interfacial and hydrodynamic parameters are involved. Hydrodynamic issues of particle dissolution have been extensively investigated both from a modeling and experimental perspective.[72]

From a microscopic (interfacial) perspective, dissolution is initiated by the solvation of molecules at the surface of the particles, which then moves to the solid–liquid interface. Then, the corresponding diffusion mechanism transfers the material to the bulk. The enthalpic component of the change in Gibb's free energy associated with solvation process arises from the average potential energy interactions whereas the entropic component is determined by the spatial conformation of the components taking part in the reaction. The breaking of the bonds in order to create the solid–liquid interface requires large enthalpic penalties. On the other hand, for the solvation of molecules requiring the extensive formation of hydrogen bonds, characterized by specific spatial arrangement, the entropic term dominates the process.[73] It is clear that the spatial orientation of the molecules in a solid particle affects the dissolution process. For example, in a crystalline particle, the exposure of functional groups promoting the formation of hydrogen bonds on the crystal planes may enhance dissolution.[74]

The Gibbs free energy per molecule at the interface between a solute molecule and the bulk solvent can be written in terms of a hard sphere model (assuming spherical solute molecules):

$$\frac{\Delta G}{4\pi R^2} \approx \frac{pR}{3} + \gamma\left(1 - \frac{2\delta}{R}\right) \tag{8.51}$$

where R is the radius of the solute, p is the pressure of the system, γ is the tension in the solute−solvent interface and δ is a length scaling parameter indicating the asymptotic behavior of the surface tension. Theoretical calculations based on this model show that for solute molecules with $R << 1$ nm, the free energy increases proportionally with the solute size. However, for $R > 1$ nm, a shift is observed, leading the free energy to grow slower, proportionally to the surface area of the solute, approaching a limiting value. This crossover marks the formation of an interface between large solute molecules and the solvent. Molecular dynamics simulations, conducted on aqueous systems, reveal that this interface formation is favored by the tendency of water molecules, held together via hydrogen bonds, to move away from the large solute molecule, to minimize the number of unformed hydrogen bonds. On average, less than one hydrogen bond is sacrificed, out of the four a water molecule can form, due to this shift. These studies verify the Stillinger's theory stating among others that[75]:

> Although liquid water might properly be described as a random, three-dimensional, hydrogen-bond network, it surely cannot have invariant fourfold coordination. Instead, some of the hydrogen bonds must be broken and others severely strained in length and direction.

8.5.3.2 *Surfactant enhanced wetting*

Surfactant-containing liquid droplets exhibits high spreading even on highly hydrophobic surfaces. The phenomenon where a liquid spreads rapidly on a surface is called superspreading.[76] The exact mechanism behind this is still unclear, nevertheless various attempts have been made to explain it using theoretical,[77] computational[78,79] and experimental tools. Despite the lack of understanding, surfactants are extensively used industrially to aid spreading, including in pharmaceutical formulations.[80]

8.5.3.3 *Effect of additives in crystallization*

The paper published by Michaels and Colville in 1960[81] sets the basis for the understanding of the role of ionic additives in crystal growth and habit. An adsorption based mechanism has been proposed to explain the changes in crystal growth. The notions of this study have been expanded for the effects of surfactants and polymers on crystal growth. Empirical and semiempirical approaches have mainly been employed, based on the functional groups and the molecular weight of the additives. However no mechanistic approach exists describing these phenomena. Recently, Kestur[82] proposed an explanation for the effects of polymers molecular weight and concentration. Nonetheless, the role of the intermolecular interactions arising from the functional groups remains unclear.

In the field of surfactant additives, apart from the changes in supersaturation driven by the increase in solubility due to the effects of surfactants, micelle formation appearing above the CMC may lead to some interesting effects. In two independent studies, one on organic[83] and the other on inorganic materials,[84] it has been shown that crystallization in solutions with surfactant concentration above the CMC, the crystal habit is significantly different. Carbamazepine dihydrate recrystallized from a solution of anhydrous carbamazepine (form III) produces elongated planar crystals with high aspect ratio. As the concentration of surfactant increases, there is a shift to a planar shape with lower aspect ratio. When the surfactant concentration is above the CMC, the shape becomes prismatic. From a theoretical perspective, the attachment energies of carbamazepine dihydrate favor the formation of the elongated shape minimizing the morphological importance of the {200} facet. This change is due to the presence of soft micellar structures.

The emergence of nonclassical nucleation[85–87] schemes could provide the tools to understand the effects of macromolecules and additives in general in crystallization. In this direction, the discovery of an intermediate liquid phase, called the *polymer-induced liquid phase*, used by organisms in biomineralization proved to be of great importance.[88,89] As discussed by Evans, features of this phenomenon can be extended in the crystallization of organic molecules with the aid of proteins via the formation of an amorphous organic phase.[90]

Coming from a different pathway, researchers have investigated the crystallization of organic compounds from the corresponding amorphous organic glasses.[91] The high energetic amorphous phase appears to enhance crystallization. In particular, even when the material is below its glass transition temperature the molecular mobility could be enough to promote nucleation and crystal growth.[92] A more intrinsic feature is that the growth rates in the bulk and the surface of the amorphous material are different, with the latter being orders of magnitude faster.[93,94] Many potential mechanisms were examined including the effects of surface tension, surface ordering, and surface mobility.[95]

Recent investigations have revealed the correlation of the growth mechanism with the fluidity at the crystal−amorphous interface.[96] Crystal growth in such a system occurs due to the expansion of the crystalline front and not because of the nucleation on the amorphous side. The crystalline side advances fast enough so to overcome diffusion phenomena arising from the fluidity of the amorphous materials. In this direction, the ratio of the diffusivity of the glassy material to the crystal growth should be less than seven picometers.

An increase in the fluidity of the amorphous material would lead to an increase of the diffusivity, resulting in crystal growth cessation.

8.6 SOLID–VAPOR INTERFACE

8.6.1 Introduction

The phenomenon of gaseous molecules attaching themselves to a solid surface can occur via two general mechanisms, physisorption and chemisorption. Of course, fundamentally this distinction is arbitrary and cases exist which are difficult to classify as being simply physisorption or chemisorption. The latter involves the formation of chemical bonds of which the enthalpy involved is comparable to the heat of reaction. In physisorption, the energy involved is similar to those required for liquefaction and is completely reversible for a nonporous solid. The third mechanism for vapor–solid interactions can also occur whereby the vapor molecules are "dissolved" into the interior of the solid, and the phenomenon is referred to as absorption. The term sorption is usually used to describe a process in which both adsorption and absorption occur. The solid is termed the adsorbent, while the vapor molecules are the adsorbate.

8.6.2 Adsorption fundamentals

Adsorption-based techniques can be employed for the determination of surface energetics of various types of materials,[97] including but not limited to polymers,[98,99] metals, and of course crystalline organic materials.[68] In this section, a brief introduction to some concepts of adsorption processes, associated with the concepts already established in the chapter, is provided. The International Union of Pure and Applied Chemistry (IUPAC) definition of adsorption is

> An increase in the concentration of a dissolved substance at the interface of a condensed and a liquid phase due to the operation of surface forces. Adsorption can also occur at the interface of a condensed and a gaseous phase.

This definition is supported mathematically by the Gibb's adsorption isotherm[100] shown in Eq. 8.52, where γ stands for the surface tension, Γ for the surface excess, μ for the chemical potential, and i for the number of components. As can be seen, this definition does not refer to solids at all.

$$d\gamma = -\sum \Gamma_i \, d\mu_i \tag{8.52}$$

In 1937, two seminal papers were published by Bangham and Radzouk, pioneering the studies of vapor–solid adsorption systems.[101,102] After that, much work on the development of new mathematical models to describe adsorption isotherms corresponding to different conditions has been undertaken. Table 8.1 summarizes some of the most well-known isotherms describing their main attributes (chronological order is followed, going from the oldest to the most recent).

8.6.3 Heterogeneous adsorption

Because of the energetic anisotropy of crystalline particles, pharmaceutical particles (as well as other materials used in engineering applications) are characterized by energetic heterogeneity. Taking into account this

TABLE 8.1 Summary of Some of the Most Important Adsorption Isotherms Available

Name	Equation	Comments
Freudlinch (1909)	$\theta_e = K_0 P_e^{\frac{1}{n}}$	Equation used on the chemisorption on activated carbon. It can describe multilayer problems
Langmuir (1918)	$\theta_e = \frac{Q_0 b P_e}{1 + b P_e}$	Well established equation suitable for chemical adsorption problems, limited to monolayer situations
BET (1938)	$\theta_{monolayer} = \frac{C P_e}{(1 - P_e)[1 + (C - 1)P_e)}$	It is used extensively for the characterization of particulate materials
Fowler-Guggenheim (1939)	$K_{FG} P_e = \frac{\theta_e}{1 - \theta_e} \exp\left(\frac{2\theta_e W}{RT}\right)$	One of the first approaches taking into account lateral interactions
Temkin (1940)	$\theta_e = \frac{RT}{\Delta Q} \ln(K_0 P_e)$	A quasi logarithmic isotherm, attempting to capture the surface heterogeneity of the material of interest
Kiselev (1958)	$K_1 P_e = \frac{\theta_e}{(1 - \theta_e)(1 + k_n \theta_e)}$	It is used extensively for the characterization of porous materials
Elovich (1962)	$\frac{q_e}{q_{monolayer}} = K_E P_e \exp\left(-\frac{q_e}{q_m}\right)$	One of the pioneer efforts for the systematic understanding of multilayer adsorption
Hill de Boer (1968)	$K_1 P_e = \frac{\theta_e}{1 - \theta_e} \exp\left(\frac{\theta_e}{1 - \theta_e} - \frac{K_2 \theta_e}{RT}\right)$	Similar to Fowler–Guggenheim, derived from van der Waals equation of state

heterogeneity parameter, the following equation has been suggested:

$$\int_0^{\infty} \chi(\varepsilon)\, d\varepsilon = 1 \tag{8.53}$$

The term $\chi(\varepsilon)$ is the probability distribution of having an energy site with energy ε on the surface of a given material. Combining the aforementioned equation with any isotherm equation, the fundamental equation of adsorption on heterogeneous surfaces is derived.[97]

$$\theta(T,p) = \int_0^{\infty} \theta_e(\varepsilon,T,p) * \chi(\varepsilon)\, d\varepsilon \tag{8.54}$$

In Eq. 8.54, θ denotes for the surface coverage and θ_e for the adsorption isotherm.

Similarly, an integral equation can be developed for the mean surface energy of adsorption:

$$\overline{E} = \int_{\overline{E}_{min}}^{\overline{E}_{max}} \varepsilon * \theta_e(\varepsilon,T,p) * \chi(\varepsilon) d\varepsilon \tag{8.55}$$

The solution of theses equations provides an insight to the energetic heterogeneity of the material under investigation. Many attempts have led to solution schemes involving different isotherm. However, there is not a universal solution scheme. Here, the attention is given to the development of a solution framework describing the adsorption phenomena on organic crystalline materials, as it takes place in IGC. The fundamentals of this effort have been set in the work of Jefferson et al.[103] and Smith et al.[104]

Jefferson's and Smith's approaches are based on the Henry's law model, implying a linear relation between surface coverage (θ) and vapor pressure. Despite the oversimplistic nature of this argument, it may be considered valid, particularly at low partial pressures. This is because pharmaceutical organic powders, at low surface coverage (<10%), will typically have an IUPAC type II isotherm, similar to BET. As can be seen, the Brunauer-Emmett-Teller (BET) equation at low pressure collapses to Henry's law.

Using the kinetic model proposed for gas adsorption on crystalline solids, Henry's constant (C) was given in the following terms:

$$C = \frac{v_0}{n_0\sqrt{2\pi mkT}} \exp\left(\frac{-\Delta G_{ads}}{kT}\right) \tag{8.56}$$

In this equation, v_0 corresponds to the mean stay time of an adsorbed molecule on the surface (ie, it measures the tendency of the molecules to desorb), n_0 stands for the concentration of surface sites, m is the mass of an adsorbed molecule, ΔG_{ads} is the adsorption surface energy, k is the Boltzmann constant, and T the temperature.

Assuming that the adsorbent molecule is a van der Waals fluid with negligibly small polar component, then the energy of adsorption per molecule is given as follows:

$$\Delta G_{ads} = -2a_m\sqrt{\gamma^{LW}_{surface}\gamma^{LW}_{adsorbate}} \tag{8.57}$$

On the surface of a crystal a finite number of patches/facets with specific surface energy value, depending on the facet, the relative coverage between two different patches is given by the ratio of the Henry's constant corresponding to each facet. In this context, it can be assumed that the concentration of surface sites is facet independent, as it is determined by the crystal lattice parameters. The adsorption of van der Waals fluids is driven only by the dispersive component of surface energy (physisorption). The mean stay time, in this case, is approximately 10^{-2} seconds. Taking into account these two assumptions, the mathematical formulation of the relative surface coverage takes the following form:

$$\frac{C_1}{C_2} = \exp\left(\frac{-\Delta G_{ads,1} + \Delta G_{ads,2}}{kT}\right) = \exp\left(\frac{2a_m\sqrt{\gamma^{LW}_{adsorbent}}\left(\sqrt{\gamma^{LW}_{surface,1}} - \sqrt{\gamma^{LW}_{surface,2}}\right)}{kT}\right) \tag{8.58}$$

8.6.3.1 *Mapping of energetic surface Heterogeneity*

Smith et al. used the analysis described previously to solve the integral equations to deconvolute representation of the dispersive surface energy profiles, assuming Gaussian distribution of surface energies. The integral equations were solved using point-by-point integration as suggested by Thielmann.[105] The underlying idea of this method is that the integral equations are solved for different combinations of the energy density distribution function ($\chi(\varepsilon)$), until an agreement is reached between model and experimental data. For example, let's assume a sample of powder χ containing crystals with four different facets $A-D$, each one with a different surface energy value (let's call it μ). Assumptions are made for the relative contribution of each of the facets to the total surface area (this fraction would be called w) of the sample. If a Gaussian distribution of the energetic sites is assumed, then the following equation for $\chi(\varepsilon)$ can be formulated:

$$\chi(\varepsilon) = \sum_{i=A}^{D} \frac{w_i}{\sigma_i\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\gamma_i^d - \mu_i}{\sigma}\right)^2} \tag{8.59}$$

In fact, since the values for μ are known from contact angle experiments, only the values of w can be varied in this equation. By inspection of the IGC data, is possible to obtain some good predictions for them.

This deterministic approach for the solution of adsorption problems is grounded on two over simplistic arguments. The first one is that Boltzmann distribution is adequate to describe, energetically, the behavior of gas molecules in a microscopic level. The second argument is that each adsorption site corresponds to a certain energetic threshold value, determined by the surface energy of the sites. If the adsorbent particles carry enough energy to overcome this threshold, then this site will be filled.

Employing arguments similar to those Einstein used to explain the stochastic nature of Brownian motion,[106] adsorption can be classified as a stochastic process. The stochastic nature of adsorption implies that even if the energy of an adsorbate does not exceed the threshold value of the adsorption site, there are still probabilities to adsorb on it. These probabilities govern both the adsorption and desorption phenomena.

8.6.4 Inverse gas chromatography (IGC)

As mentioned, particulate materials are energetically heterogeneous. Thus, adsorption methods can be used to characterize their surface properties since these techniques can provide an idea for the values of surface energy at different values of surface coverage. Inverse gas chromatography is one of those methods. The application of this method in the field of the particulate materials has started in the mid-1970s.[107–110] The name of this process is indicative of its fundamental features. It is exactly the inverse of a conventional gas chromatography, developed by the Russian botanist Mikhail Tswett[111] at the beginning of the 20th century. A known gas probe is injected into a column packed with the material of interest. The time needed for the gas to exit the column (retention time) is used to determine the net retention volume, which in turn is used to determine a material's physicochemical properties, including the surface energy.

The infinite dilution regime of IGC, corresponds to the situation where a small amount of probe is injected into the system covering only a small amount of the surface of the solid. The back tailing regime corresponds to relatively high surface coverage ($>25\%$) of macroscopic materials, the isotherm of which corresponds to type II and IV IUPAC isotherms. The fronting situation is one corresponding to either type III or V IUPAC isotherms.[112] These isotherms characterize materials with low surface energy. This situation is usually undesirable. Low surface energy materials do not adsorb the injected probes well. Desorption phenomena prevail, and usually the results are not reliable. In cases like this, the best option is to run the experiment at a lower temperature until the fronting effects are minimized.

From the chromatograms obtained, the retention time of each injected probe can be estimated. If a Gaussian chromatogram is produced, then the retention time is taken as the time at which the maximum peak occurs. If the peak is skewed, then the center of mass should be taken. The retention volume can be estimated from the following equation:

$$V_N = \frac{j}{m_s} w(t_R - t_0)\left(\frac{T_s}{T_{Ref}}\right) \tag{8.60}$$

The retention time (t_R) can be implemented in the equation given here to calculate the elution volume (V_N). In the same equation, j stands for the James–Martin pressure drop coefficient. This parameter acts as a correction factor accounting for the compressibility of the injected probes. Parameter m_s is the specific surface area of the material under investigation in m^2 per unit mass, w is the carrier gas flow rate in standard cubic centimeters per minute, t_0 is the dead time, the time takes for an inert probe molecule to pass through the column, T_s is the temperature at which the experiment is performed, and T_{Ref} is the reference temperature, which is usually taken to be 273.15 K.

Taking into account the aspects outlined in the surface energy section of this chapter, it is clear that the selection of the probes is important in this method. Usually, linear alkanes are used for the determination of the dispersive component of the surface energy. Due to the differences in adsorption behavior of each of them, a series of n-alkanes is used (usually C6-C11). The results can be analyzed based either on the Schultz[113,114] or the Dorris and Gray[115] methods, to calculate the dispersive surface energy of the sample. The results obtained from the two methods are similar, and there is not a rational way to select which is most appropriate for each case.[115,116]

Similarly for the determination of the polar component of the surface energy, polar probes are used.[117] The experimental results from different temperatures can be plotted based on the fundamental thermodynamic equation shown here:

$$\Delta G_{ij}^0 = \Delta H_{ij}^0 - T\Delta S_{ij}^0 \tag{8.61}$$

To determine the enthalpy of adsorption for the acid-base components, experiments should be performed as a function of temperatures. More recently, the use of monopolar probes such as ethyl acetate and dichloromethane have been proposed by Das et al. to determine the acid-base properties by IGC.[118]

8.6.5 Implications of solid–vapor interfaces

8.6.5.1 Moisture content in solid-state materials

Solid-state formulations are exposed in high moisture environment during processing and/or storage. Water in crystalline solids can be present as bound or unbound water. The term bound water indicates that the water molecules are incorporated into the crystal lattice or in channels. On the other hand, unbound water is either adsorbed on the crystal facets or trapped in voids. This distinction is useful in distinguishing moisture containing pharmaceutical formulations, but can be misleading.[119,120]

For crystalline materials, when adsorption is dictated by van der Waals forces (physisorption), then multiple layers of adsorbate molecules can be formed. In this context, BET and Hill de Boer adsorption isotherms are particularly attractive for use in sorption studies. On the other hand, when chemical (short-range) forces dominate, then monolayer adsorption prevails where the Langmuir isotherm is adequate. It is important to appreciate that because of the nature of the interactions, physisorption is a slower process, requiring a longer time for equilibrium to be established. Experience has shown that the validity of these equations is confined in a certain range. Nonetheless, more complicated equations which may provide a larger validity range may require careful determination of the associated parameters.

In amorphous solids, above the T_g, the adsorbed moisture diffuses in the bulk of the material. In this case, the system is treated in terms of liquid–liquid mixing models (Flory–Huggins)[121,122] or models describing the solvent diffusion in amorphous polymers (Vrentas–Vrentas).[123] In any case, it should be mentioned that moisture absorption in amorphous solids is not uniform. It has been shown that, depending on the conditions, solvent clusters can be formed in the solid. Of course, the use of these models is not limited to moisture sorption studies, but it has been extensively used in the characterization of polymer-drug interactions.[124–126]

The hysteresis behavior of the adsorption isotherms observed in sorption studies provides information on the amorphous content and the porosity of the material under investigation, apart from the details on the water uptake capacity. Generally speaking, hysteresis can occur due to the porosity of the material, but also on structural defects and chemical impurities on the surface of the material of interest. In Fig. 8.8, some examples of hysteresis behavior are shown. Crystalline macroporous materials exhibit no or marginal hysteresis; porous materials have a larger hysteresis and amorphous materials lead to an open loop hysteresis. The shape of the hysteresis loop is dependent on the size and geometry of the pores and is categorized by the IUPAC classification. In amorphous materials, the open loop indicates that the water bound in the bulk of the material cannot be released with just reversing the RH.

8.6.5.2 Drying

Moisture removal includes simultaneous heat and mass transfer; energy is transferred from the surroundings to the particle leading to the evaporation of the moisture. Due to the different states at which moisture is bound in the solid, drying processes exhibit different stages. During the initial heating period, the fluid is heated to evaporation. Then there is a constant drying rate phase the moisture adsorbed on the surfaces is removed until a critical moisture content value corresponding to the tightly bound water content. The vapor pressure of the moisture bound in capillary structures can be significantly lower due to the effects of confinement described by the Kelvin equation.

In the case of crystalline solids, the exposure of crystal planes to ambient moisture after drying may induce

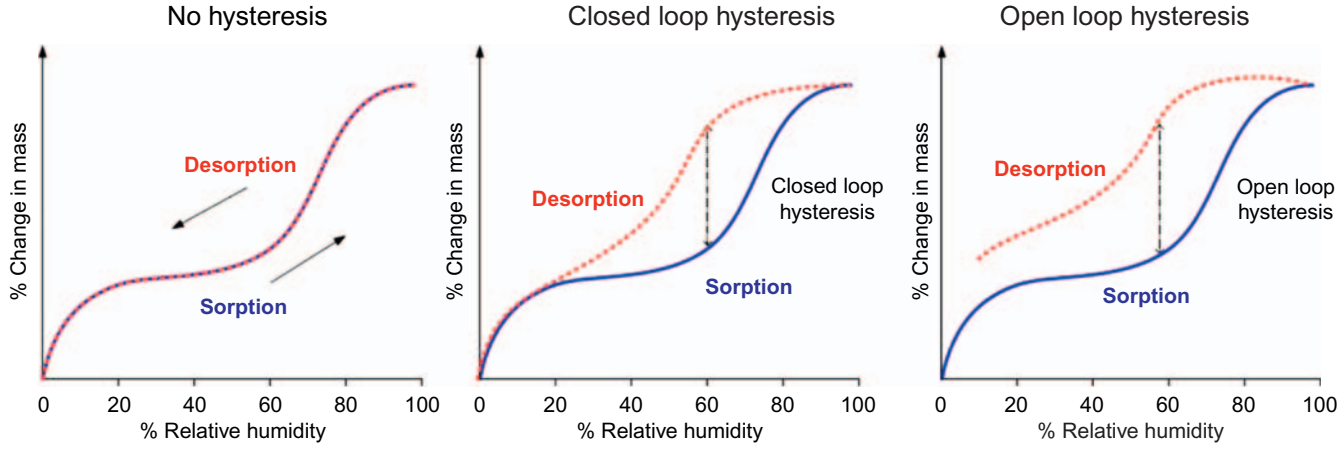


FIGURE 8.8 Sorption desorption isotherms for different hysteresis cases.

either changes on the surface patterning or increase in nucleation density. The increase in surface roughness associated with the latter is a common feature of hydrated compounds. An example obtained from the dehydration of carbamazepine dihydrate is shown in Fig. 8.9. However, no literature findings support the facet-specific dependence of this process, although crystal defects can affect this process.[127–129]

Crystal defects can be distinguished in three different categories[130]:

1. Point defects: this case includes situations when atoms or molecules are literally missing from a point of the crystal lattice, creating a vacant point, or when an extra atom appears on a random position on the lattice.
2. Linear defects: this is when a group of atoms/molecules is misplaced in the lattice creating screw or edge dislocations.
3. Planar defects: in this case defects appear in the region between two homogeneous crystal planes

Defects can arise randomly during crystallization or can be induced by thermal or mechanical stress. They are characterized by higher energy than the bulk of the material, and thus, they are sites of preferential reaction (either via Lifshitz–van der Waals or chemical interactions). When hydrates are dried under vacuum, the nucleus of the anhydrous form is initially formed in the area around the crack (Fig. 8.10). The moisture removed from the material provides the molecular mobility on the interface between anhydrous and hydrated material to facilitate the anhydrous material growth. If the moisture removal is too fast (eg, under vacuum conditions), then the crystal lattice cannot withstand the stress associated with the removal of bound water molecules and collapses to an amorphous state.[131] The amorphous, nonequilibrium, material seeks to recrystallize. However, recrystallization is not necessarily taking place toward the most stable anhydrous polymorph, but toward the one with the energy minimum of closest proximity.[132] Thus, it could turn to an intermediate form which would then turn to the most stable one under the effect of Ostwald's rule, which can be summarized by the following excerpt from Threlfall[133]:

> At a sufficiently high supersaturation, the first form that crystallises is the most soluble form. This transient state then transforms to the more stable form through a process of dissolution and crystallization.

Apart from the conventional drying approaches, spray drying and freeze drying allow moisture removal combined with the transition from crystalline phase to amorphous. This transition directly affects the molecular arrangement in the bulk and the surface of the solid, since amorphous solids are characterized by a lack of thermodynamic equilibrium. Existing studies, are limited in the use of infinite dilution inverse gas chromatography for the determination of the surface energetics of spray and freeze-dried materials, meaning that they are taking into account only a negligibly small amount of surface area of the material under investigation. The data obtained from these studies show negligible changes in the surface energy. Nonetheless, further studies should be conducted on the ground of finite dilution inverse gas chromatography to provide us with positive results for the effects of spray and freeze drying on the surface energetics of solids.[134]

The dynamic behavior of suspensions in a spray dryer can lead to the formation of particles with a wide range of characteristics. The mechanism for the formation of these hollow particles is determined by the heat and mass transfer processes described by the Peclet number. Suspensions are injected into the spray drying chamber and under the proper conditions the habit of the dried particles, at the outlet of the spray

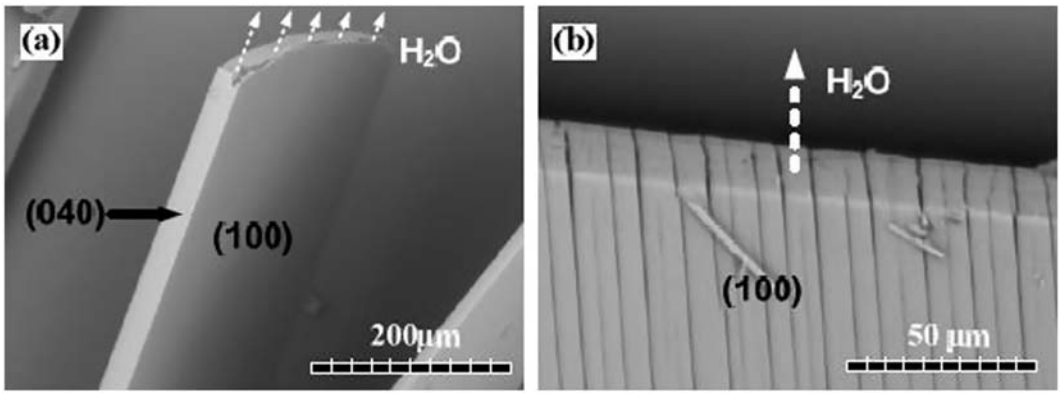


FIGURE 8.9 The effects of drying on the surface roughness of carbamazepine dihydrate, the cracks are induced by the removal of bound water.[127]

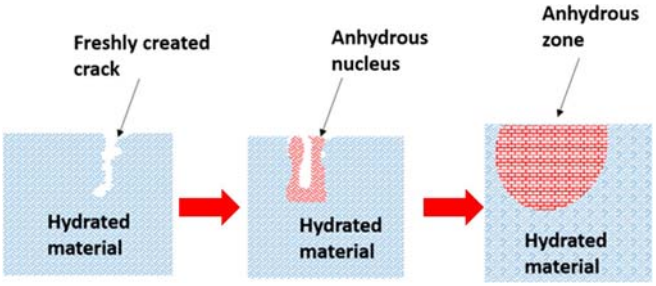


FIGURE 8.10 The development of an anhydrous form during dehydration at the cracks of the surface of a hydrated crystalline solid.

dryer, can be engineered. The dimensionless Peclet number, providing the ratio of the change in surface area of a particle containing bubble with the diffusion coefficient of the solids in the liquid droplet can be modified with the use of coated particles with different diffusion coefficients.

In his works on spray drying, Vehring prove the applicability of this process in the manufacturing of hollow particles to be used in drug delivery systems. The formation of hollow particles can also lead to the decrease in the bulk density of the product; though can be susceptible to breakage. The versatility of spray drying opens the possibility for particle engineering and product design, including pharmaceutical solids.[135,136]

8.7 INTERFACIAL PHENOMENON (SOLID–SOLID)

8.7.1 Fundamental thermodynamics

When solid particles interact, the behavior of the resultant system is greatly influenced by their interfacial interparticulate interactions. This interaction may be defined by adhesion or cohesion forces. The attractive force between particles of different compounds is regarded as adhesion while the attractive force between particles of the same compound is termed as cohesion. Adhesion and cohesion forces are the sum total of van der Waals forces, capillary forces, mechanical interlocking and solid bridging in addition to minor contributions from contact potential and Coulombic forces. Van der Waals forces are considered to be the principal forces for particulate adhesion or cohesion that include interactions like dipole-dipole, dipole-induced dipole and induced dipole-induced dipole. Adhesion forces have been discussed in the section on fundamental intermolecular forces, however, they are briefly mentioned here in the context of powdered solids.

The relative contribution of these forces is mainly determined by material properties such as hydrophilicity or hydrophobicity, particle size and local environmental conditions, such as RH.[137] These factors are particularly important for DPIs, as micronized powders are highly susceptible to interparticulate forces because of their high specific surface area.

Surfaces of pharmaceutical powder particles are often rough. Particle interactions can thus be reduced by surface roughness since the latter increases the distance between particles and consequently decreases contact and van der Waals forces. However, if the surface is excessively rough, for example, crevices and holes are present on the surface, they may interlock which can in turn dramatically increase interparticular interactions by mechanical interlocking. In addition, if a small particle is trapped in a depression on a large particle, the attractive forces are more than the required detachment forces due to intermolecular and electrostatic interactions.[138]

The presence of moisture has a significant effect on interparticulate interactions. Although van der Waals forces exist in both low and high humidity, low level of moisture can enhance the dissipation of triboelectric charge and decrease interparticulate interactions. At high RH, water vapor moving through interparticulate spaces can condense and form liquid bridges and influence interparticulate forces by at least three different mechanisms such as by changing surface energy, surface conductivity through electron mobilization and surface tension.[139] The presence of high level of moisture can develop capillary forces and increase interparticulate interactions.

The development of capillary forces and their contribution to overall adhesion forces is governed by several factors. Firstly, capillary forces are not significant at an RH of less than 50%, and they are the dominant contributor to adhesion at an RH of above 65%.[140] Secondly, capillary forces develop more easily for hydrophilic compounds than hydrophobic compounds. Liquid bridges can be formed by the moisture condensed between particles. If the lifetime of these bridges is high, they can start dissolving the particles. Due to evaporation of the moisture, the dissolved solid is recrystallized to form solid bridges. Solid bridges increase the interparticulate interactions significantly.

Solid bridges can also be formed by mineral bridges between particles of the same material; chemical interaction or reaction between particles; partial melting of low melting point solids caused by friction or heat and subsequent cooling. Although solid bridging is rare in practice, they become important during long-term storage of hydrophilic compounds.[141] Liquid bridges are often followed by solid bridges due to recrystallization.[142]

The surface energy of particles plays an important role in their behavior. The term surface energy is sometimes confused with adhesion/cohesion. In the previous section, it has been clearly shown how surface energy is related with adhesion/cohesion forces. Adhesion force can be calculated from surface energy by using contact mechanics models. The history of these models dates back to 1881 when Hertz[143] published his influential work accurately describing the deformations on macroscopic systems, but not taking in account the effect of adhesion forces. Numerous advancements have been made, and different models have been proposed since then and concentrated especially on the effects of adhesion in smaller scales.[144–146] The two most well-known models are the Derjaguin-Muller-Toporov (DMT) model[25] and Johnson-Kendall-Roberts (JKR) model.[147]

In the JKR model, the forces outside the contact area between the two bodies are considered negligible. In addition, at the edges of the contact area, the elastic stress tends to infinity meaning that the strength of the system increases unboundedly with decreasing size of the sphere.[148,149] On the other hand, in DMT model, the surrounding forces appear to have an effect on the adhesion, nevertheless they do not affect the shape of a sphere which is deformed according to Hertz's model. This discussion can simplistically be summarized by the statement that JKR model is more appropriate for larger systems where the sphere has a large enough contact surface area to make the effect of the surrounding forces negligible, and smaller systems are better treated through DMT model.[150] In this context, the more frequent question is what can be considered being large and what being small. Maugis attempts to answer this question in his work. Using a more general model developed by Dugdale, it proves that the aforementioned models are specific cases of Dugdale's model and has related the two models with dimensionless parameters.[148]

8.7.2 Experimental techniques

For the quantitative determination of adhesion force, atomic force microscope (AFM) is widely used. Inverse gas chromatography can be used to calculate adhesion/cohesion forces by measuring surface energy. However, scanning electron microscope (SEM) and other imaging techniques are also used to understand the interparticulate interactions qualitatively.

8.7.2.1 Atomic force microscope

AFM has been extremely useful for measuring adhesion force. Prior to the introduction of AFM, optical microscope was used to image a surface. The AFM is a type of scanning probe microscope that has a resolution in the order of nanometer. It has revolutionized the imaging of particulate surfaces, thus has been useful in manipulating particulate surface in the nanometer range. The AFM has also been used to determine adhesion forces and surface energy from DMT/JKR models.[151,152]

The cantilever is made of a well-characterized material, and its Hooke's law (spring) constant is known. The tip of the cantilever is of known shape, and its material properties are also known. A small force slowly pushes the tip of the cantilever toward the surface of the material. The movement of the cantilever from the moment it approaches the surface enough to feel the attractive forces until it gets pushed back is used to calculate the force via Hooke's law. The force-to-distance behavior is displayed in Fig. 8.11 in conjunction with the cantilever surface relationship.

Initially, the tip approaches the surface when is in a distance where the attractive forces start to become significant, it bends. As it bends, the effect of attractive forces becomes more profound until the tip attaches to the surface. Then the tip is pulled back slowly until it is totally dispatched from the surface.

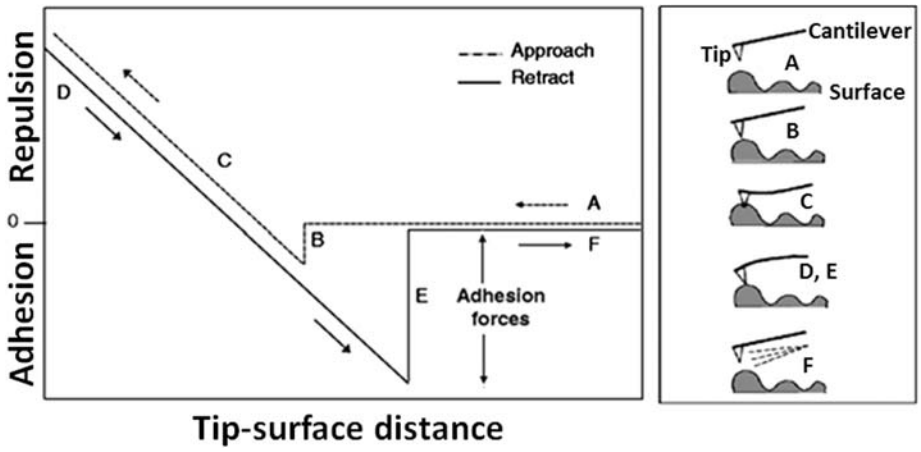


FIGURE 8.11 Schematic showing the force curve during an AFM experiment in conjunction with cantilever surface relation.

8.7.2.2 Scanning electron microscope

SEM helps the qualitative identification of drug-carrier interactions on the surface of the particle. In normal SEM, a vacuum is necessary, which was not possible to see the changes made by different environmental humidity on powders. The introduction of environmental SEM in the 1980s has made now possible to study powder sample controlling temperature and RH.[153]

8.7.3 Pharmaceutical implications

Solid dosage forms dominate all the other dosage forms, with oral solid dosage forms constituting over 80% of the market. Solid dosage forms include tablets, capsules, suppositories and dry powder for reconstitution or inhalation. In many phases of their production process, understandings of the interfacial interparticulate interactions are important. The processes include but not limited to milling, flowability, mixing, tabletting, etc. The influence of solid–solid interfacial interactions will be discussed in the following section.

8.7.3.1 Flowability

Flowability is an important bulk powder characteristic. The term "Flowable" means an irreversible deformation of a powder to make it flow due to the application of external energy or force. Various parameters such as angle of repose, Carr's index, Hausner ratio, flow function (ff) are used to express flowability of powders.

The angle of repose is the simplest method to express the flowability of powders. The angle of the heaped cone of a free standing powder is the angle of repose. When a powder is allowed to fall on a flat surface from a funnel positioned at a certain height, the funnel is gradually moved upward in order to maintain a fixed height between the powder tip and the bottom of the funnel. The angle produced by the powder on the surface is the angle of repose. The other way is to allow powder from a central hole of a flat-bottomed container. The slope produced by the remaining powder in the container bottom is the angle of repose. Furthermore, rotating drum can be used to measure the angle of repose. The material is placed, homogeneously, in the drum and left to rotate for a number of revolutions. Then the angle is measured.

Another refined method to measure flowability is the ff, defined as the ratio of consolidation stress (σ_1) to unconfined yield strength or uniaxial compression stress (σ_c).

The flow of powder particles is influenced by interparticulate interactions which include adhesion/cohesion, frictional forces, mechanical forces due to interlocking. A range of other factors can also influence flowability of powders. For example, particle size and size distribution, particle shape and shape distribution, porosity, density, moisture content, surface composition, substances present in powder.

The lower the adhesion/cohesion forces, the greater will be the flowability. Flowability can be improved by using both hydrophilic such as aerosol and hydrophobic substances such as magnesium stearate, sodium stearate, magnesium silicate and calcium silicate, etc. Hydrophobic additives can increase flowability of powders by decreasing interparticulate interactions by reducing van der Waals forces at low RH. In contrast, hydrophilic additives may decrease flowability by increasing interparticulate interactions by developing capillary forces. The use of low cohesive, hydrophobic lubricants has, already, been employed to improve powder's flowability.

One of the challenges of increasing the flowability by using flow additives is to disperse the flow additives on the surface of host particles since very fine particles of flow additives might prefer to agglomerate themselves. Flow additives are added either by mixing or by mechanofusion. The processes of mixing and mechanofusion is discussed in the following.

8.7.3.2 Mixing or blending

Mixing or blending is a widely used process in pharmaceutical industries to blend a multicomponent particular system. In the formulation of tablets, active ingredients are mixed with excipients before granulation. Different glidants and lubricants are mixed with granules after granulations and before tabletting. Similarly, drugs are mixed with various excipients for other solid dosage forms such as capsules, dry powders for inhalation and powders for reconstitution. The mixing is carried out using different mixers or blenders, for example, V-mixer, Cone mixer, rotating cylinder, etc. It is expected that a mixture will be homogeneous meaning the different components will be distributed in a uniform manner throughout the system.

If drug particles are small in size, they are cohesive and can form agglomerates. The mixing mechanisms involve two stages: the breakdown of agglomerates to primary particles and the dispersion of primary particles. An ideal mixing process may break down these agglomerates to primary particles and disperse throughout the system in a manner that they are evenly distributed, and it will be maintained for a long time after mixing so that the ultimate dosage forms such as tablets and capsules have uniformity in drug content. In practice, it is not possible to have a completely uniform (ideal) mixture. Often, a "satisfactory level of random mixing" is achieved, where the

average (mean) of each type of particles in a series of samples is close to the ideal mean and the variability in drug content is low.

There are three main procedures of solid mixing: diffusive mixing, convective mixing, and shear mixing. In diffusive mixing, individual particles move to change position relative to neighboring particles in a random way. This type of mixing is possible for free-flowing powders and is facilitated by the differences in density and particle size. In convective mixing, a portion of the mixture is taken from one place and deposited in another. Finally, in shear mixing, particles are forced to move past each other. In this type of mixing, usually large agglomerates are broken up but not small agglomerates. This is determined by the relative strength of agglomerates and the shear applied.

Demixing (segregation) can occur at the same time as mixing occurs when the external mixing procedure stops. The factors promoting segregation include differences in particle size, shape and density. If the particles are uniform in size, shape and density, and the surface is smooth, achieving uniformity of mixture is still extremely difficult when the drug concentration is very low in the mixture, and the drug or excipients are highly cohesive resulting in very strong agglomerates of drug or excipients.[154,155]

In order to improve the homogeneity of a powder mixture containing a very low concentration of a drug, ordered mixing process was introduced by John Hersey.[156] This technique was initially developed for producing tablets of highly potent drugs. Later on, it has been regularly used to produce carrier based DPIs where only a small percent of drug is mixed with large carriers. Each "ordered unit" of a mix will be similar or almost identical in composition to all other ordered units in the mixture. Ordered mixing is achieved by dividing and recombining the powder bed in different ways such as preblending and geometric dilution; "doubling."

When the drug is mixed with carriers, the homogeneity ultimately depends on the balance between cohesiveness of drugs and excipients and the adhesion between drugs and excipients. It is expected that adhesion between drug and excipients will be higher than cohesive forces of drugs.

Measuring the surface energy of powder mixtures, using inverse gas chromatography, is not a straightforward process. The differences between physically mixed and coated samples can be identified. Following the discussion conducted in the solid–vapor interface section of this chapter, it is easy to appreciate that inverse gas chromatography study of mixtures comprising materials with significantly different surface energy involves certain limitations arising from the nature of the technique. The fluid probes interact preferentially with the material exhibiting higher energy adsorption sites on the basis of a Boltzmann distribution, leaving the adsorption sites of the lower energy material uncovered. Thus, especially at low values of surface coverage, only one of the materials may take place in the measurement.

The surface energy of a powder after mixing might be different due to the creation of new surfaces due to attrition/fracturing.[157] Recently, surface energy profile determination explained the mixing phenomenon even better. The nonpolar surface energy profile of the salbutamol sulfate (SS)–magnesium stearate (MgSt) mixture to be positioned between those of SS and MgSt and it was within 5% range of nonpolar surface energy profiles of MgSt, indicating that the surface of SS was covered by MgSt during the mixing process.[158] Similarly, the surface energy profiles of two blends of fine lactose (LH210) and large lactose (LH250) at different ratios were found to be positioned between those of LH210 and LH250.[159]

Interparticulate interactions were also studied by using AFM to understand the influence of mixing two or more components on the mixture especially in the context of DPI performance. It was concluded that one of the mechanisms on how the inclusion of fines improves performance of drugs from a career based formulation is that when fines are mixed with large carriers, the highly active sites are occupied by the fines, making the less active sites available for drugs that can be easily dispersed.[160]

8.7.3.3 *High-shear mixing or dry coating*

Mechanofusion is a process of mechanical dry coating in which submicron-sized guest particles directly attach onto relatively larger, micron-sized host or core particles. This process uses shear and compression to produce a thin continuous film of the guest material on the surface of the host. This high-shear mixing process has been used to change the surface properties of powders by coating a material with hydrophobic materials to modify the surface energy and reduce particle interactions.[161]

Dry coating can be performed by using a magnetic assisted impaction coater (MAIC) or a hybridizer (HB)[162], or by using a mechanofusion apparatus.[163] Magnesium stearate, traditionally used as a lubricant, has been widely used in coating by mechanofusion. Dry coating of powders such as fine lactose, salbutamol sulfate, triamcinolone acetonide, and salmeterol xinafoate with a low cohesive material such as magnesium stearate has been very successful at increasing the deagglomeration of those powders.[7,164] Mechanofusion of lactose with MgSt has been shown to improve flow properties and dispensability dramatically compared with conventional mixing of MgSt with lactose in a turbular mixer.[165]

Other materials include hydrophobic and hydrophilic silica. Coating of cornstarch with different silicas using MAIC has shown to improve flowability. Hydrophobic silica was found to be more effective than hydrophilic silica, especially by reducing the liquid bridge formation.[162]

Recently, the Resonant Acoustic Mixer was developed to improve mixing and dispersion processes. It applies resonance phenomenon to effectively transfer energy to the content in the mixing container. In the case of mixing particulate materials, the powders are accelerated as high as 100 times of the gravitation force. Numerous particle-particle and particle-vessel collisions lead to the effectiveness of the process. Its application in dry powder coating of pharmaceuticals was demonstrated by Mullarney and coworkers using several drugs.[166] Flexibility in mixing scale (from subgram to hundreds of kilogram), the ability to contain the mixing contents, and the availability of the continuous version of the mixer make this technology a great prospect for applications in pharmaceutical development.

8.7.3.4 Milling

Micronization and milling are commonly used pharmaceutical processes for reducing particle size of pharmaceutical materials.[167–171] A reduction in size is extremely important in the case of DPI since particles of $<5\,\mu m$ size can reach lower respiratory tracts which are the primary sites of drug delivery through inhalation. In general, micronization of large drug particles is used to produce particles of an inhalable size and then mixed with a large carrier in a carrier-based formulation for inhalation. It is also important to increase the dissolution of solids since the reduction in particle size can improve surface area, and thereby, the dissolution rate of the materials. In tablet manufacturing, drug is micronized before wet granulation with excipients. Micronization can be carried out using high-energy air jet mill while ball milling is used. The process is influenced by many factors, such as the nature of starting material, temperature, RH and energy input, etc.

Micronization introduces new interfaces by fracturing particle, creating lattice defects, or by producing amorphous materials. These new interfaces may have different exposed surface groups than the surface of the original material. These processes, therefore, may increase or decrease the surface energy of materials based on the exposed surface groups. In general, an increase in dispersive or nonpolar surface energy was observed upon micronization of compounds, such as paracetamol Form I,[109] salbutamol sulfate,[172] salmeterol xinafoate, and crystalline α-lactose monohydrate.[173] Increase in surface energy may result in increased agglomeration due to their cohesive forces exceeding the gravitation forces, thus, creating new challenges in drug delivery. Milling of DL-propranolol initially increased dispersive surface energy until a particular point, brittle to ductile the transition point, followed by a decrease with further milling.[174] Similar change was observed for ketoconazole and griseofulvin.[175]

However, micronization of compounds showed mixed effects (increase or decrease) on polar surface energy. For example, the acidic component of polar surface energy was found to decrease for DL-propranolol,[174] salbutamol sulfate,[172] salmeterol xinafoate and α-lactose monohydrate,[176] but it increased for acetaminophen upon micronization. The basic component of surface energy was increased for DL-propranolol,[177] salbutamol sulfate,[172] and α-lactose monohydrate,[176] but it decreased for acetaminophen upon micronization or milling. Micronization and milling may expose new surfaces, created by the formation of lattice defects or because of the formation of amorphous sites.[172,178–180] In all the aforementioned studies, surface energy was determined at infinite dilution using an inverse gas chromatography and results may only provide a limited description of powder properties as such measurements do not fully describe the heterogeneity of crystalline solids.

The particle surface energy characteristics were revealed in more detail with the determination of surface energy (heterogeneity) using finite dilution IGC. The distribution of nonpolar surface energy was more heterogeneous for milled lactose than untreated α-lactose monohydrate.[67] The milled lactose had significantly higher nonpolar surface energy than untreated (before milling) α-lactose monohydrate at infinite dilution (zero surface coverage). However, at higher surface coverage, the milled lactose had lower surface energy than the untreated α-lactose monohydrate. In another study, it was observed that, compared to ibipinabant before micronization, the nonpolar surface energy of ibipinabant after micronization was smaller at low surface coverages whereas it was larger at higher surface coverages.[181] The increase in nonpolar surface energy was attributed to the generation and redistribution of high energy sites on micronization.

8.7.3.5 Tableting

The ability of a powder to form a tablet of adequate mechanical strength is termed as tabletting. It is important that a tablet remain intact during its manufacturing, packaging and transportation processes until it is consumed by the patient. It is also important that the tablet is not excessively strong so that it does not undergo disintegration and dissolution after oral administration. The mechanical strength of a tablet is expressed as the tensile strength which is the minimum force required to break a tablet.[182,183] The tensile

strength of a tablet depends on many characteristics of the consolidating granules or powders, such as particle size, shape, density, surface area and interparticulate interactions.[184]

A tablet is produced by compression of granules or powders of drugs with or without excipients. When a powder is compressed, it undergoes processes such as particle rearrangement, particle fracture, and particle deformation. Interparticulate interactions play a major role in tabletting. Interparticulate interactions such as van der Waals forces, hydrogen bonding may exist between molecular solids while ion-dipole, ion-induced dipole or Coulombic forces may exist for ions. During the early compression process, powder particles come closer due to rearrangement of particles. The interparticulate interactions such as van der Waals forces increase when the distance between particles decreases. As the compression force is increased, the particles fracture (depending on the brittle-ductile nature and the strength of the particles). Again the strength of an individual particle depends on the interparticulate interactions of the molecules and ions they are made of. The last phase of consolidation by compression is particle deformation which involves plastic and elastic deformation with an increase in interparticulate interactions by increasing particle contact area. The ultimate success of tabletting depends on interparticulate interactions of powder particles, material properties, and compression process.

8.7.3.6 Triboelectrification

When two materials undergo friction, the particles may be charged or electrified due to the transfer of electrons from one particle to the other; this charging is called *triboelectric charging*, *tribo-charging*, or *frictional electrification*.[185] *Contact charging* denotes the charging which arises due to the transfer of electrons when two different materials come in contact and are separated. Charging due to short contact for collision is called *impact charging*. It is practically difficult to differentiate the purposes of charge transfer; therefore, *triboelectric charging* is commonly used to refer to all three of these processes.[186]

Triboelectric charging could be generated due to the direct transfer of electrons between two materials in contact due to the difference in their work functions (as in metal-metal contact) or surface charges. It may also arise due to the transfer of ions or materials from surfaces during contact, frictions or impact. For a metal-metal contact, the amount of charge transferred (ΔQ_c) can be calculated as the product of the contact potential difference (V_c) and the capacitance (C_0) between the two bodies.

Since the particles collide with each other, or with the wall of the container, or particles are abraded during production and handling of powdered solids (eg, milling, micronization, sieving, mixing, coating and spray drying), triboelectric charge generation in the particles is common.[187] This can be both problematic and beneficial. The negative impact of triboelectrification of particles includes particle adhesion during manufacturing, particle accumulation on equipment surface, segregation and altered dose uniformity.[188–192] In contrast, the positive impact of electrostatic charging includes the production of ordered mixtures of charged powder particles exploiting the opposite polarity.[193] This type of ordered mixture is more stable and may improve content homogeneity more than other powder processing techniques.[190,193,194] Factors that influence triboelectric charge production include particle size, shape, surface roughness,[195,196] surface chemistry, presence of functional groups and chemical structure[197–199] and electrical properties of powders and contacting surfaces.[200,201]

Triboelectric charge is affected by the presence of moisture (ie, environmental RH and temperature), nature of contacting material, and velocity of particles.[186] At a specific temperature, the triboelectric charge decreases with increasing RH probably due to increased leakage for decreased electrical resistance on the surface.[202] It also decreases with increasing temperature.

In order to minimize triboelectric charging, particle surface is coated with different materials. The irradiation using ultraviolet rays and plasma processing are the two methods commonly used to change surface charge of polymers. In electrophotography, two types of materials, charge control agents and charge stabilizers are used to minimize charging. Fumed silica and polyvinylidene fluoride are used as surface charge control agents while amines or quaternary ammonium salts are bulk charge control agents which are blended with polymers to control charging. Polyester salts are used as charge stabilizers.

Triboelectric charges can be characterized by different techniques. One of these techniques is a cascade method.[203] The triboelectric charges are produced on particles when they are allowed to fall down from the top of a reference plate which is kept tilted at a certain angle. The charge is measured by a Faraday cage. Higashiyama et al.[204] described a slightly different technique. In their technique, particles are allowed to fall from one end of a charging plate to the other end through vibration where the charge of the particles is measured by a Faraday cage. Luga et al.[205] used a fluidized bed to develop triboelectric charges and a charge spectrometer to measure charge distribution.

8.8 FUTURE DIRECTIONS—OPINIONS

The interfacial properties and behavior of solid-state pharmaceutical formulations are complex, though

ubiquitous in the manufacture, handling, formulation and delivery of pharmaceuticals. However, much of the current descriptors employ semiempirical approaches, and there is a need in providing the framework for more fundamental understanding. The ability to develop linkages between quantum and nanoscale phenomena with their manifestations in the microscale will lead to major scientific advances and have significant industrial relevance. Molecular modeling tools and population balance equations could play a significant role in bridging these gaps. Development of theoretical models across lengthscales; from molecules to particles to process,[206,207] are currently undertaken, some of which are already resulting in more accurate descriptions and useful insights.

Many fundamental equations consider the ideal case where particles are spherical and the surfaces are flat and smooth. In reality, such ideal cases do not exist, and surface morphology such as surface roughness and shape should be considered in describing interfacial phenomena. It is plausible that the inclusion of the anisotropic nature of crystalline solids[208,209] into current theoretical models will leads to a better understand and ultimately predictability of powder processability. In parallel, advancement in characterization techniques is also necessary. Approaches to describe heterogeneity or distribution of particle properties are crucial, as single average value do not accurately explain powder behavior. For example, advances in determining surface energy heterogeneity, crystal size and crystal shape distribution of powders have been made over the last decade.

Recently developed imaging techniques allow the determination of both the size and shape distribution of particles.[210,211] Significant steps have been made toward the fundamental understanding of the mechanisms leading to the crystallization of particles with different crystal habits under different conditions.[212–214] The surface energy deconvolution of powders is also a useful tool in the understanding of the effects of surface anisotropy. However, these advancements have not matured enough to be implemented in the existing mechanistic models describing macroscopic phenomena like aggregation and attrition.

Whilst not addressed exhaustively in this chapter, novel formulation methods including but not limited to amorphous solid dispersions, nanodispersions, porous nanoparticles and liposomes exhibit interesting interfacial phenomena. Since some of them rely on particles in the submicron scale and/or self-assembly interactions it is, also, obvious that interfacial phenomena would probably dictate the corresponding product/process development into a great extent. Thus, the question is not if we would need to go deeper in understanding the implications of interfaces in pharmaceutical formulations but if we would be able to progress fast enough to reach the demands for more advanced formulations. We are in a more advantageous position compared with the pioneers in the field of interfaces. Novel experimental techniques enable us to go achieve a level of accuracy that our predecessors did not enjoy. Also, we have tools to develop and solve more sophisticated mathematical models to describe interfacial phenomena, than anyone could imagine back in Gibbs's era, in the late 19th century. Nonetheless, interfaces remain a challenging field where significant work should be done to bridge different length scales and provide integrated understanding for their behavior.

References

1. Gibbs JW. *Graphical methods in the thermodynamics of fluids*vol. 1 *The collected works of J. Willard Gibbs*. 1st ed New York, NY: Longmans, Green and Co; 1906. p. 1–32.
2. Trasatti S, Parsons R. Interphases in systems of conducting phases: recommendations 1985 Supersedes provisional version published 1983. *J Electroanal Chem Interfacial Electrochem* 1986;**205**(1–2):359–76.
3. Mei R, Shang H, Klausner JF, Kallman E. A contact model for the effect of particle coating on improving the flowability of cohesive powders. *KONA Powder Part J* 1997;**15**:132–41.
4. Tomas J. Fundamentals of cohesive powder consolidation and flow. *Gran Mat* 2004;**6**:75–86.
5. Bridgwater J. Fundamental powder mixing mechanisms. *Powder Technol* 1976;**15**(2):215–36.
6. Chaudhuri B, Mehrotra A, Muzzio FJ, Tomassone MS. Cohesive effects in powder mixing in a tumbling blender. *Powder Technol* 2006;**165**(2):105–14.
7. Zhou QT, Qu L, Larson I, Stewart PJ, Morton DAV. Improving aerosolization of drug powders by reducing powder intrinsic cohesion via a mechanical dry coating approach. *Int J Pharm* 2010;**394**(1–2):50–9.
8. Shah UV, Olusanmi D, Narang AS, Hussain MA, Tobyn MJ, Heng JYY. Decoupling the contribution of dispersive and acid-base components of surface energy on the cohesion of pharmaceutical powders. *Int J Pharm* 2014;**475**(1–2):592–6.
9. Shah UV, Olusanmi D, Narang AS, Hussain MA, Tobyn MJ, Hinder SJ, Heng JYY. Decoupling the contribution of surface energy and surface area on the cohesion of pharmaceutical -powders. *Pharm Res* 2015;**32**(1):248–59.
10. Han X, Ghoroi C, To D, Chen Y, Davé R. Simultaneous micronization and surface modification for improvement of flow and dissolution of drug particles. *Int J Pharm* 2011;**415**(1–2):185–95.
11. Giry K, Péan JM, Giraud L, Marsas S, Rolland H, Wüthrich P. Drug/lactose co-micronization by jet milling to improve aerosolization properties of a powder for inhalation. *Int J Pharm* 2006;**321**(1–2):162–6.
12. Ho R, Naderi M, Heng JYY, Williams DR, Thielmann F, Bouza P, Keith AR, Thiele G, Burnett DJ. Effect of milling on particle shape and surface energy heterogeneity of needle-shaped crystals. *Pharm Res* 2012;**29**:2806–16.
13. Ho R, Dilworth SE, Williams DR, Heng JYY. Role of surface chemistry and energetics in high shear wet granulation. *Ind Eng Chem Res* 2011;**50**:9642–9.
14. Costa P, Sousa Lobo JM. Modeling and comparison of dissolution profiles. *Eur J Pharm Sci* 2001;**13**(2):123–33.

15. Burnett DJ, Thielmann F, Booth J. Determining the critical relative humidity for moisture-induced phase transitions. *Int J Pharm* 2004;**287**(1–2):123–33.
16. Burnett DJ, Thielmann F, Sokoloski T, Brum J. Investigating the moisture-induced crystallization kinetics of spray-dried lactose. *Int J Pharm* 2006;**313**(1–2):23–8.
17. Khoo JY, Williams DR, Heng JYY. Dehydration kinetics of pharmaceutical hydrate: effects of environmental conditions and crystal forms. *Drying Techol* 2010;**28**:1164–9.
18. Khoo JY, Shah UV, Schaepertoens M, Williams DR, Heng JYY. Process-induced phase transformation of carbamazepine dihydrate to its polymorphic anhydrates. *Powder Technol* 2013;**236**: 114–21.
19. van der Waals JD. *The equation of state for gases and liquids*. *Nobel lectures, physics 1901-1921*. 1st ed Amsterdam: Elsevier Publishing Company; 1967. p. 254–65.
20. Tabor D. Surface forces and surface interactions. *J Colloid Interf Sci* 1977;**58**(1):2–13.
21. Israelachvili JN. *Intermolecular and surface forces*. 3rd ed Oxford: Elsevier Academic Press; 2011.
22. Lifshitz EM. The theory of molecular attractive forces between solids. *Sov Phys* 1956;**2**(1):73–83.
23. Rytov S. *Theory of electric fluctuations and thermal radiations*. 1st ed Moscow: Moscow Academy of Sciences Press; 1959.
24. van Oss JC, Chaudhury KM, Good RJ. Interfacial Lifshitz-van der Waals and polar interactions in macroscopic systems. *Chem Rev* 1988;**88**:927–41.
25. Derjaguin BV, Landau L. Theory of the stability of strongly charged lyophobic sols and of the adhesion of strongly charged particles in solutions of electrolytes. *Prog Surf Sci* 1941;**43** (1):30–59.
26. Verweym EJW, Overbeek JTG. *Theory of the stability of lyophobic colloids*. Amsterdam: Elsevier Publishing Company Inc; 1948.
27. Fischer EW. Electronenmikroskopische Untersuchungen zur Stabilitat von Suspensionen in makromolekularen Losungen. *Kol Zeit* 1958;**160**(2):120–41.
28. Fowkes FM. Dispersion force contributions to surface and interfacial tensions, contact angles, and heats of immersion. In: Gould RF, editor. *Contact angle, wettability and adhesion*. 1st ed Washington, DC: American Chemical Society; 1964. . p. 99–111.
29. Owens DK, Wendt RC. Estimation of the surface free energy of polymers. *J Appl Polym Sci* 1969;**13**:1741–7.
30. London F. Zur theorie und systematik der molekularkräfte. *Zeit Phys* 1930;**63**(3-4):245–79.
31. Girifalco FA, Good RJ. A theory for the estimation of surface and interfacial energies. I. Derivation and application to interfacial tension. *J Phys Chem* 1957;**61**(7):904–9.
32. Wu S. Calculation of interfacial tension in polymer systems. *J Polym Sci* 1971;**34**(1):1741–7.
33. Fowkes FM, Mostafa AM. Acid-base interactions in polymer adsorption. *Ind Eng Chem Res* 1978;**17**(1):3–7.
34. Drago RS, Vogel CG, Needham ET. A four-parameter equation for predicting enthalpies of adduct formation. *J Am Chem Soc* 1971;**93**(23):6014–26.
35. Mulliken RS. Lewis acids and bases and molecular complex. *J Chem Phys* 1971;**93**:6014–26.
36. Gutmann V. Coordination chemistry of certain transition-metal ions. The role of solvent. *Coord Chem Rev* 1966;**2**:239–56.
37. Fahr A, Liu X. Drug delivery strategies for poorly water-soluble drugs. *Exp Op Drug Del* 2007;**4**(4):403–16.
38. Pouton CW. Formulation of poorly water-soluble drugs for oral administration: physicochemical and physiological issues and the lipid formulation classification system. *Eur J Pharm Sci* 2006;**29**(3–4):278–87.
39. Young T. An essay on the cohesion of fluids. *Phil Trans R Soc Lond* 1805;**95**:65–87.
40. Good RJ. A thermodynamic derivation of Wenzel's modification of Young's equation for contact angles; together with a theory of hysteresis. *J Am Chem Soc* 1952;**74**(20):5041–2.
41. de Gennes PG. Wetting: statics and dynamics. *Rev Mod Phys* 1985;**57**:827–63.
42. de Gennes PG, Brochard-Wyart F, Quere D. *Capillarity and wetting phenomena*. 1st ed. New York, NY: Springer; 2004.
43. Bormashenko E, Pogreb R, Whyman G, Erlich M. Resonance Cassie-Wenzel wetting transition for horizontally vibrated drops deposited on a rough surface. *Langmuir* 2007;**23**:12217–21.
44. Bahadur V, Garimella VS. Preventing the Cassie-Wenzel transition using surfaces with noncommunicating roughness elements. *Langmuir* 2009;**25**(8):4815–20.
45. Bormashenko E, Bormashenko Y, Stein T, Whyman G, Bormashenko E. Why do pigeon feathers repel water? Hydrophobicity of pennae, Cassie–Baxter wetting hypothesis and Cassie–Wenzel capillarity-induced wetting transition. *J Colloid Interf Sci* 2007;**311**:212–16.
46. Nosovonsky M, Bhusan B. Biomimetic superhydrophobic surfaces: multiscale approach. *Nano Let* 2007;**7**(9):2633–7.
47. Tadmor R. Line energy and the relation between advancing, receding, and Young contact angles. *Langmuir* 2004;**20**(18):7659–64.
48. Li D, Neumann AW. Equation of state for interfacial tensions of solid-liquid systems. *Adv Colloid Interf Sci* 1992;**39**:299–345.
49. Neumann AW, Good RJ, Hope CJ, Sejpal M. An equation-of-state approach to determine surface tensions of low-energy solids from contact angles. *J Colloid Interf Sci* 1974;**49**(2):291–304.
50. Heng JYY. *Anisotropic surface properties of crystalline pharmaceutical solids (PhD)*. Imperial College London; 2006.
51. Wilhelmy L. Ueber die abhangigkeit der capillaritats–constanten des walkohols von substanz und gestalt des benetztenjesten korpers. *Ann Phys Chim* 1863;**119**:179–217.
52. Dove JW, Buckton G, Doherty C. A comparison of two contact angle measurement methods and inverse gas chromatography to assess the surface energies of theophylline and caffeine. *Int J Pharm* 1996;**138**:199–206.
53. Bartell FE, Osterhof HJ. Determination of the wettability of a solid by a liquid. *Ind Eng Chem* 1927;**19**:1277–80.
54. Brakel JV, Heertjes PM. Capillary rise in porous media. Part I: a problem. *Powder Technol* 1977;**16**:75–81.
55. Brakel JV, Heertjes PM. Capillary rise in porous media. Part II: secondary phenomena. *Powder Technol* 1977;**16**:83–90.
56. Brakel JV, Heertjes PM. Capillary rise in porous media. Part III: role of the contact angle. *Powder Technol* 1977;**16**:91–6.
57. Cassie ABD. Wettability of porous surfaces. *Trans Faraday Soc* 1944;**40**:546–50.
58. Cassie ABD. Contact angles. *Dis Faraday Soc* 1948;**3**:11–16.
59. Wenzel RN. Resistance of solid surfaces to wetting by water. *Ind Eng Chem* 1936;**28**:988–94.
60. Burnett DJ, Heng JYY, Thielmann F, Garcia RA, Naderi M, Acharya M. Measuring surface roughness of pharmaceutical powders using vapor sorption methods. *AAPS Pharm Sci Tech* 2011;**12**(1):56–61.
61. Anvir D, Farin D, Pfeifer P. Molecular fractal surfaces. *Nature* 1984;**308**:261–3.
62. Pfeifer P. Chemistry in noninteger dimensions between two and three. I. Fractal theory of heterogeneous surfaces. *J Chem Phys* 1987;**79**(7):3558–65.
63. Begat P, Morton DAV, Staniforth NJ, Price R. The cohesive-adhesive balances in dry powder inhaler formulations I: direct quantification by Atomic Force Microscopy. *Pharm Res* 2004;**21** (9):1591–7.

64. Begat P, Morton DAV, Staniforth JN, Price R. The cohesive-adhesive balances in dry powder inhaler formulations II: influence on fine particle delivery characteristics. *Pharm Res* 2004;**21**(10):1826–33.
65. Beach ER, Tormoen GW, Drelich J, Han R. Pull-off force measurements between rough surfaces by Atomic Force Microscopy. *J Colloid Interf Sci* 2002;**247**(1):84–99.
66. Janus J, Fauxpoint G, Arntz Y, Pelletier H, Etienne O. Surface roughness and morphology of three nanocomposites after two different polishing treatments by a multitechnique approach. *Dent Mat* 2010;**26**(5):416–25.
67. Thielmann F, Burnett DJ, Heng JYY. Determination of the surface energy distributions of different processed lactose. *Drug Del Ind Pharm* 2007;**33**(11):1240–53.
68. Ho R, Heng JYY. A review of inverse gas chromatography and its development as a tool to characterize anisotropic surface properties of pharmaceutical solids. *KONA Powder Part J* 2013;**30**:164–80.
69. Karde V, Ghoroi C. Influence of surface modification on wettability and surface energy characteristics of pharmaceutical excipient powders. *Int J Pharm* 2014;**475**(1–2):351–63.
70. Karde V, Ghoroi C. Fine powder flow under humid environmental conditions from the perspective of surface energy. *Int J Pharm* 2015;**485**(1–2):192–201.
71. Karde V, Panda S, Ghoroi C. Surface modification to improve powder bulk behavior under humid conditions. *Powder Technol* 2015;**278**:181–8.
72. Dressman J, Kramer J. *Pharmaceutical dissolution testing*. 1st ed. Boca Raton, FL: Taylor & Francis; 2005.
73. Chandler D. Interfaces and the driving force of hydrophobic assembly. *Nature* 2005;**437**:640–7.
74. Prasad KVR, Ristic RI, Sheen DB, Sherwood JN. Dissolution kinetics of paracetamol single crystals. *Int J Pharm* 2002;**238**(1–2):29–41.
75. Stillinger HF. Structure in aqueous solutions of nonpolar solutes from the standpoint of scaled-particle theory. *J Sol Chem* 1973;**2**(2/3):141–58.
76. Hill RM. Superspreading. *Curr Op Colloid Interf Sci* 1998;**3**(3):247–54.
77. Nikolov A, Wasan D. Superspreading mechanisms: an overview. *Eur Ph J* 2011;**197**:325–41.
78. Theodorakis PE, Müller EA, Craster RV, Matar OK. Insights into surfactant-assisted superspreading. *Curr Op Colloid Interf Sci* 2014;**19**(4):283–9.
79. Theodorakis PE, Muller AE, Craster VR, Matar OK. Superspreading: mechanisms and molecular design. *Langmuir* 2015;**31**(8):2304–9.
80. Tadros TF. *Surfactants in pharmaceutical formulations*. *Applied surfactants*. 1st ed. Germany: Wiley-VCH; 2005. p. 433–502.
81. Michaels AS, Colville AR. The effect of surface active agent on crystal growth rate and crystal habit. *J Phys Chem* 1960;**64**(1):13–19.
82. Kestur US, Lee H, Santiago D, Rinaldi C, Won YY, Taylor LS. Effects of the molecular weight and concentration of polymer additives, and temperature on the melt crystallization kinetics of a small drug molecule. *Cryst Growth Des* 2010;**10**:3585–95.
83. Rodriguez-Hornedo N, Murphy D. Surfactant-facilitated crystallization of dihydrate carbamazepine during dissolution of anhydrous polymorph. *J Pharm Sci* 2004;**93**(2):449–60.
84. Singh N, Yeri R, Chakraborty J. Effect of ionic surfactants and alcohols on the morphology of $CuSO4 \cdot 5H2O$ crystals: combined use of factors and significance of threshold surfactant concentration. *Ind Eng Chem Res* 2013;**52**(43):15041–8.
85. Piana S, Jones F, Gale JD. Assisted desolvation as a key kinetic step for crystal growth. *J Am Chem Soc* 2006;**128**(41):13568–74.
86. Gerbauer D, Volkel A, Colfen H. Stable prenucleation calcium carbonate clusters. *Science* 2008;**320**(5909):1819–22.
87. Cartwright HEJ, Checa AG, Gale JD, Gerbauer D, Sainz-Diaz I. Calcium carbonate polyamorphism and its role in biomineralization: how many amorphous calcium carbonates are there? *Angew Chem* 2012;**51**(48):11960–70.
88. Gower LB, Odom DJ. Deposition of calcium carbonate films by a polymer-induced liquid-precursor (PILP) process. *J Cryst Growth* 2000;**210**(4):719–34.
89. Bewernitz M, Gebauer M, Long J, Colfen H, Gower LB. A metastable liquid precursor phase of calcium carbonate and its interactions with polyaspartate. *Faraday Dis* 2012;**159**:291–312.
90. Evans JS. "Liquid-like" biomineralization protein assemblies: a key to the regulation of non-classical nucleation. *CrystEngComm* 2013;**15**(42):8388–94.
91. Hikima T, Adachi Y, Hanaya M, Oguni M. Determination of potentially homogeneous-nucleation-based crystallization in o-terphenyl and an interpretation of the nucleation-enhancement mechanism. *Phys Rev B* 1995;**52**(6):3900–9.
92. Yoshioka M, Hancock BC, Zografi G. Crystallization of indomethacin from the amorphous state below and above its glass transition temperature. *J Pharm Sci* 1994;**83**(12):1700–5.
93. Zanotto ED. Experimental studies of surface nucleation and crystallization of glasses. *Ceram Trans Nucl Crystal Liquids Glasses* 1993;**30**:65–74.
94. Zhu L, Wong L, Yu L. Surface-enhanced crystallization of amorphous nifedipine. *Mol Pharm* 2008;**5**(6):921–6.
95. Sun Y, Zhu L, Kearns KL, Ediger MD, Yu L. Glasses crystallize rapidly at free surfaces by growing crystals upward. *Proc Natl Acad Sci* 2011;**108**(15):5990–5.
96. Musumeci D, Powell CT, Ediger MD, Yu L. Termination of solid-state crystal growth in molecular glasses by fluidity. *Phys Chem Lett* 2014;**5**:1705–10.
97. Rudzinski W, Everett DH. *Adsorption of gases on heterogeneous surfaces*. San Diego, CA: Academic Press; 1992.
98. Chehimi MM, Abel M, Perruchot C, Delamar M, Lascelles SF, Armes SP. The determination of the surface energy of conducting polymers by inverse gas chromatography at infinite dilution. *Synthetic Met* 1999;**104**(1):51–9.
99. Zhao L, Choi P. Determination of solvent-independent polymer–polymer interaction parameter by an improved inverse gas chromatographic approach. *Polymer* 2001;**42**(3):1075–81.
100. Gibbs JW. *Collected works*. New York, NY: Longmans, Green; 1928.
101. Bangham DH, Razouk RI. Adsorption and the wettability of solid surfaces. *Trans Faraday Soc* 1937;**33**:1459–63.
102. Bangham DH. Adsorption and the wettability of solid surfaces. *Trans Faraday Soc* 1937;**33**:805–11.
103. Jefferson AE, Williams DR, Heng JYY. Computing the surface energy distributions of heterogeneous crystalline powders. *J Adhes Sci Technol* 2011;**25**(4-5):339–55.
104. Smith RR, Williams DR, Burnett DJ, Heng JYY. A new method to determine dispersive surface energy site distributions by Inverse Gas Chromatography. *Langmuir* 2014;**30**(27):8029–35.
105. Thielmann F. Introduction into the characterisation of porous materials by inverse gas chromatography. *J Chromatogr A* 2004;**1037**(1–2):115–23.
106. Einstein A. Investigations on the theory of Brownian movement. *Annal Phys* 1905;**17**:549–60.
107. Braun J, Guillet JE. Determination of crystallinity of olefin copolymers and polyolefin powders by inverse gas chromatography. *Polym Chem* 1975;**13**(5):1119–31.
108. Hudec P. Determination of crystallinity of polyethylene powder by inverse chromatography. *Macromol Chem Phys* 1977;**78**(4):1187–93.

109. Heng JYY, Thielmann F, Williams DR. The effects of milling on the surface properties of form I paracetamol crystals. *Pharm Res* 2006;**23**(8):1918–27.
110. Yla-Maihaniemi PP, Heng JYY, Thielmann F, Williams DR. Inverse gas chromatographic method for measuring the dispersive surface energy distribution for particulates. *Langmuir* 2008;**24**(17):9551–7.
111. Tswett MS. Physical chemical studies on chlorophyll adsorptions. *Berich Deutsch Botan Gesel* 1906;**24**:316–23.
112. Balbuena BP, Gubbins EK. Theoretical interpretation of adsorption behavior of simple fluids in slit pores. *Langmuir* 1993;**9**(7):1801–14.
113. Hamieh T, Schultz J. New approach to characterise physicochemical properties of solid substrates by inverse gas chromatography at infinite dilution: I. Some new methods to determine the surface areas of some molecules adsorbed on solid surfaces. *J Chromatogr A* 2002;**969**(1–2):17–25.
114. Hamieh T, Schultz J. New approach to characterise physicochemical properties of solid substrates by inverse gas chromatography at infinite dilution: II. Study of the transition temperatures of poly(methyl methacrylate) at various tacticities and of poly(methyl methacrylate) adsorbed on alumina and silica. *J Chromatogr A* 2002;**969**(1–2):27–36.
115. Dorris GM, Gray DG. Adsorption of n-alkanes at zero surface coverage on cellulose paper and wood fibers. *J Colloid Interf Sci* 1980;**77**(2):353–62.
116. Shi B, Wang Y, Jia L. Comparison of Dorris–Gray and Schultz methods for the calculation of surface dispersive free energy by inverse gas chromatography. *J Chromatogr A* 2011;**1218**(6):860–2.
117. Panzer U, Schreiber PH. On the evaluation of surface interactions by Inverse Gas Chromatography. *Macromolecules* 1992;**25**:3633–7.
118. Das CS, Larson I, Morton DAV, Stewart PJ. Determination of the polar and total surface energy distributions of particulates by Inverse Gas Chromatography. *Langmuir* 2011;**27**(2):521–3.
119. Authelin J. Thermodynamics of non-stoichiometric pharmaceutical hydrates. *Int J Pharm* 2005;**303**(1–2):37–53.
120. Febles M, Perez-Hernandez N, Perez C, Rodriguez LM, Foces-Foces C, Roux MV, Morales EQ, Buntkowsky G, Limbach HH, Martin JD. Distinct dynamic behaviors of water molecules in hydrated pores. *J Am Chem Soc* 2006;**128**(31):10008–9.
121. Flory PJ. Thermodynamics of high polymer solutions. *J Chem Phys* 1941;**9**(8):660–1.
122. Huggins LM. Solutions of long chain compounds. *J Chem Phys* 1941;**9**(5):440.
123. Vrentas JS, Vrentas CM. Sorption in glassy polymers. *Macromolecules* 1991;**24**(9):2404–12.
124. Taylor LS, Zografi G. Spectroscopic characterization of interactions between PVP and indomethacin in amorphous molecular dispersions. *Pharm Res* 1997;**14**(12):1691–8.
125. Shamblin SL, Taylor LS, Zografi G. Mixing behavior of colyophilized binary systems. *J Pharm Sci* 1998;**87**(6):694–701.
126. Taylor LS, Zografi G. Fourier Transform Raman spectroscopic study of the interaction of water vapor with amorphous polymers. *J Pharm Sci* 2000;**90**(7):888–901.
127. Khoo JY. *Thermodynamic stability and kinetic analysis of pharmaceutical channel hydrate during dehydration process (PhD Thesis)*. London: Imperial College; 2010.
128. Hutternauch R, Fricke S. The influence of lattice defects on the course and rate of drying of particulate solids. *Int J Pharm Tech Prod Manufactur* 1981;**2**:35–7.
129. Galwey AK. The topochemistry of the reaction interface in decomposition processes in solids. *React Solid* 1990;**8**(3–4):211–30.
130. Kelly A, Groves GW, Kidd P. *Crystallography and crystal defects*. 2nd ed. Chichester: John Wiley & Sons Ltd; 2000.
131. Li Y, Han J, Zhang GGZ, Grant JW, Suryanarayanan R. In situ dehydration of carbamazepine dihydrate: a novel technique to prepare amorphous anhydrous carbamazepine. *Pharm Dev Technol* 2000;**5**(2):257–66.
132. Threlfall T. Structural and thermodynamic explanations of Ostwald's rule. *Org Process Res Dev* 2003;**7**:1017–27.
133. Ostwald W. Studdien über die bildung und umwandlung fester körper. *Zeit Phys Chem* 1897;**22**:289–330.
134. Peng Y, Gardner DJ, Han Y, Cai Z, Tshabalala MA. Influence of drying method on the surface energy of cellulose nanofibrils determined by inverse gas chromatography. *J Colloid Interf Sci* 2013;**405**:85–95.
135. Vehring R, Foss WR, Lechuga-Ballesteros D. Particle formation in spray drying. *J Aerosol Sci* 2007;**38**(7):728–46.
136. Vehring R. Pharmaceutical particle engineering via spray drying. *Pharm Res* 2008;**25**(5).
137. Coelho MC, Harnby N. Moisture bonding in powders. *Powder Technol* 1978;**20**(2):201–5.
138. Zeng XM, Martin GP, Marriott C, Pritchard J. The effects of carrier size and morphology on the dispersion of salbutamol sulphate after aerosolization at different flow rates. *J Pharm Pharmacol* 2000;**52**:1211–21.
139. Zeng XM, Martin GP, Marriott C, Pritchard J. The influence of carrier morphology on drug delivery by dry powder inhalers. *Int J Pharm* 2000;**200**(1):93–106.
140. Zimon AD. *Adhesion in a gas medium. Sources of adhesion. Adhesion of dust powder*. New York, NY: Consultants Bureau; 1982. p. 93–144.
141. Stewart P. Particle interaction in pharmaceutical systems. *Pharm Inter* 1986;**7**:146–9.
142. Padmadisastra Y, Kennedy RA, Stewart PJ. Solid bridge formation in sulphonamide-Emdex interactive systems. *Int J Pharm* 1994;**112**(1):55–63.
143. Hertz VH. Ueber die Beruhrung fester elastischer Korper. *J Reine Angew Math* 1881;**92**:156–71.
144. Muller VM, Yushchenko VS, Derjaguin BV. On the influence of molecular forces on the deformation of an elastic sphere and its sticking to a rigid plane. *J Colloid Interf Sci* 1980;**77**(1):91–101.
145. Muller VM, Yushchenko VS, Derjaguin BV. General theoretical consideration of the influence of surface forces on contact deformations and the reciprocal adhesion of elastic spherical particles. *J Colloid Interf Sci* 1983;**92**(1):92–101.
146. Shi X, Zhao PY. Comparison of various adhesion contact theories and the influence of dimensionless load parameter. *J Adhes Sci Tech* 2004;**18**(1):55–68.
147. Johnson KL, Kendall K, Roberts AD. Surface energy and the contact of elastic solids. *P Roy Soc Lond A Mat* 1971;**324**(1558):301–13.
148. Muller VM, Derjaguin BV, Toporov YP. On two methods of calculation of the force of sticking of an elastic sphere to a rigid plane. *Colloids Surf* 1983;**7**(3):251–9.
149. Gao H, Yao H, Ji B. Nanostructures in biological materials. In: Hannik RHJ, Hill AJ, editors. *Nanostructure control of materials*. 1st ed. Cambridge: Woodhead Publishing; 2006. p. 115–46.
150. Maugis D. Adhesion of spheres: the JKR-DMT transition using a dugdale model. *J Colloid Interf Sci* 1992;**150**(1):243–69.
151. Binnig G, Quate CF, Gerber C. Atomic force microscope. *Phys Rev Lett* 1986;**56**(9):930.
152. Binnig G, Gerber C, Stoll E, Albrecht TR, Quate CF. Atomic resolution with Atomic Force Microscope. *Europhys Lett* 1987;**3**(12):1281–6.
153. Danilatos GD. Foundations of environmental scanning electron microscopy. *Adv Electronics Electr Phys* 1988;**18**:662–6.
154. Salter GF, Farnish RJ, Bradley MSA, Burnett AJ. Segregation of binary mixtures of particles during the filling of a

two-dimensional representation of a hopper. *J Proc Mech Eng* 2000;**214**(3):197–208.
155. Deng T, Paul KA, Bradley MSA, Immins L, Preston C, Scott JF, et al. Investigations on air induced segregation of pharmaceutical powders and effect of material flow functions. *Powder Technol* 2010;**203**(2):354–8.
156. Hersey JA. Ordered mixing: a new concept in powder mixing practice. *Powder Technol* 1975;**11**(1):41–4.
157. Buckton G, Choularton A, Beezer AE, Chatham SM. The effect of the comminution technique on the surface energy of a powder. *Int J Pharm* 1988;**47**(1–3):121–8.
158. Tay T, Das S, Stewart P. Magnesium stearate increases salbutamol sulphate dispersion: what is the mechanism? *Int J Pharm* 2010;**383**(1–2):62–9.
159. Ho R, Muresan AS, Hebbink GA, Heng JYY. Influence of fines on the surface energy heterogeneity of lactose for pulmonary drug delivery. *Int J Pharm* 2010;**388**(1–2):88–94.
160. Jones MD, Price R. The Influence of fine excipient particles on the performance of carrier-based dry powder inhalation formulations. *Pharm Res* 2006;**23**(8):1665–74.
161. Pfeffer R, Dave RN, Wei D, Ramlakhan M. Synthesis of engineered particulates with tailored properties using dry particle coating. *Powder Technol* 2001;**117**(1–2):40–67.
162. Yang J, Sliva A, Banerjee A, Dave RN, Pfeffer R. Dry particle coating for improving the flowability of cohesive powders. *Powder Technol* 2005;**158**(1–3):21–33.
163. Das SC, Zhou Q, Morton DAV, Larson I, Stewart PJ. Use of surface energy distributions by inverse gas chromatography to understand mechanofusion processing and functionality of lactose coated with magnesium stearate. *Eur J Pharm Sci* 2011;**43**(4):325–33.
164. Zhou QT, Armstrong B, Larson I, Stewart PJ, Morton DAV. Understanding the influence of powder flowability, fluidization and de-agglomeration characteristics on the aerosolization of pharmaceutical model powders. *Eur J Pharm Sci* 2010;**40**(5):412–21.
165. Das CS, Stewart PJ. Characterising surface energy of pharmaceutical powders by inverse gas chromatography at finite dilution. *J Pharm Pharmacol* 2012;**64**(9):1337–48.
166. Polizzi M, Langdon BA, Beach LE, Mullarney MP. Applying dry powder coatings. *Pharm Technol* 2011;**35**(10):94–102.
167. Kwan CC, Mio H, Qi Chen Y, Ding Y, Saito F, Papadopoulos DG, Bentham AC, Ghadiri M. Analysis of the milling rate of pharmaceutical powders using the Distinct Element Method (DEM). *Chem Eng Sci* 2005;**60**(5):1441–8.
168. Ghadiri M, Kwan CC, Ding Y. Analysis of milling and the role of feed properties. In: Salman AD, Ghadiri M, Hounslow MJ (Ed.). *Handbook of Powder Technology*. Amsterdam: Elsevier Science B.V. 2007. p. 605-634.
169. Kwan CC, Chen YQ, Ding Y, Papadopoulos DG, Bentham AC, Ghadiri M. Development of a novel approach towards predicting the milling behaviour of pharmaceutical powders. *Eur J Pharm Sci* 2004;**23**(4–5):327–36.
170. Ning Z, Ghadiri M. Distinct element analysis of attrition of granular solids under shear deformation. *Chem Eng Sci* 2006;**61**(18):5991–6001.
171. Ghadiri M, Zhang Z. Impact attrition of particulate solids Part 1. A theoretical model of chipping. *Chem Eng Sci*. 2002;**57**:3659–69.
172. Feeley JC, York P, Sumby BS, Dicks H. Determination of surface properties and flow characteristics of salbutamol sulphate, before and after micronisation. *Int J Pharm* 1998;**172**(1–2):89–96.
173. Newell HE, Buckton G. Inverse gas chromatography: investigating whether the technique preferentially probes high energy sites for mixtures of crystalline and amorphous lactose. *Pharm Res* 2004;**21**:1440–4.
174. York P, Ticehurst MD, Osborn J, Roberts R, Rowe R. Characterization of surface energetics of milled DL propranolol hydrochloride using inverse gas chromatography and molecular modeling. *Int J Pharm* 1998;**174**:1793–804.
175. Otte A, Carvajal MT. Assessment of milling-induced disorder of two pharmaceutical compounds. *J Pharm Sci* 2010;**100**(5):1793–804.
176. Grimsey I, Feeley P, York P. Analysis of the surface energy of pharmaceutical powders by inverse gas chromatography. *J Pharm Sci* 2002;**91**:571–83.
177. Ticehurst MD, Basford PA, Dallman CI, Lukas TM, Marshall PV, Nichols G, Smith D. Characterisation of the influence of micronisation on the crystallinity and physical stability of revatropate hydrobromide. *Int J Pharm* 2000;**193**(2):247–59.
178. Florence A, Salole E. Changes in crystallinity and solubility on comminution of digoxin and observations on spironolactone. *J Pharm Pharmacol* 1976;**248**:479–80.
179. Egawa H, Maeda S, Yonemochi E, Oguchi T, Yamamoto K, Nakai Y. Solubility parameter and dissolution behaviour of cefalexin powders with different crystallinity. *Chem Pharm Bull* 1992;**40**:819–20.
180. Briggner L, Buckton G, Bystrom K, Darcy P. The use of isothermal microcalorimetry in the study of changes in crystallinity induced during the processing of powders. *Int J Pharm* 1994;**105**(2):125–35.
181. Gamble JF, Leane M, Olusanmi D, Tobyn M, Šupuk E, Khoo J, Naderi M. Surface energy analysis as a tool to probe the surface energy characteristics of micronized materials—a comparison with inverse gas chromatography. *Int J Pharm* 2012;**422**(1–2):238–44.
182. Sinka C. Modelling powder compaction. *KONA Powder Part J* 2007;**25**:4–22.
183. Sinka IC, Pitt KG, Cocks ACF. The strength of pharmaceutical tablets. In: Salman AD, Ghadiri M, Hounslow MJ (Ed.). *Handbook of Powder Technology*. Amsterdam: Elsevier Science B.V. 2007. p. 941–970.
184. Sun CC. Decoding powder tabletability: roles of particle adhesion and plasticity. *J Adhes Sci Tech* 2001;**25**(4-5):483–99.
185. Matsusaka S, Masuda H. Electrostatics of particles. *Adv Powder Technol* 2003;**14**(2):143–66.
186. Matsusaka S, Maruyama H, Matsuyama T, Ghadiri M. Triboelectric charging of powders: a review. *Chem Eng Sci* 2010;**65**(22):5781–807.
187. Ghori MU, Šupuk E, Conway BR. Tribo-electric charging and adhesion of cellulose ethers and their mixtures with flurbiprofen. *Eur J Pharm Sci* 2014;**65**:1–8.
188. Hussain T, Kaialy W, Deng T, Bradley MSA, Nokhodchi A, Armour-Chélu D. A novel sensing technique for measurement of magnitude and polarity of electrostatic charge distribution across individual particles. *Int J Pharm* 2013;**441**(1–2):781–9.
189. Pu Y, Mazumder M, Cooney C. Effects of electrostatic charging on pharmaceutical powder blending homogeneity. *J Pharm Sci* 2008;**98**(7):2412–21.
190. Staniforth JN, Rees JE. Powder mixing by triboelectrification. *Powder Technol* 1981;**30**:255–6.
191. Staniforth JN, Rees JE. Electrostatic charge interactions in ordered powder mixes. *J Pharm Pharmacol* 1982;**34**:69–76.
192. Supuk E, Hassanpour A, Ahmadian H, Ghadiri M, Matsuyama T. Triboelectrification and associated segregation of pharmaceutical bulk powders. *KONA Powder Part J* 2011;**29**:208–23.
193. Maki R, Suihko E, Rost S, Heiskanen M, Murtomaa M, Lehto VP, Ketolainen J. Modifying drug release and tablet properties of

starch acetate tablets by dry powder agglomeration. *J Pharm Sci* 2007;**96**:438–47.
194. Karner S, Urbanetz NA. The impact of electrostatic charge in pharmaceutical powders with specific focus on inhalation-powders. *J Aerosol Sci* 2011;**42**(6):428–45.
195. Carter PA, Cassidy OE, Rowley G, Merrifield DR. Triboelectrification of fractionated crystalline and spray-dried lactose. *Pharm Pharmacol Comm* 2011;**4**(2):111–15.
196. Eilbeck J, Rowley G, Carter PA, Fletcher EJ. Effect of materials of construction of pharmaceutical processing equipment and drug delivery devices on the triboelectrification of size-fractionated Lactose. *Pharm Pharmacol Comm* 1999;**5**(7):429–33.
197. Kamiyama M, Maeda M, Okutani H, Koyama K, Matsuda H, Sano Y. Effect of functional groups on the triboelectric charging property of polymer particles. *J Appl Polym Sci* 1994;**51**(9): 1667–71.
198. Mazumder MK, Sims RA, Biris AS, Srirama PK, Saini D, Yurteri CU, Trigwell S, De S, Sharma R. Twenty-first century research needs in electrostatic processes applied to industry and medicine. *Chem Eng Sci* 2006;**61**(7):2192–211.
199. Shinohara I, Yamamoto F, Anzai H, Endo S. Chemical structure and electrostatic properties of polymers. *J Electrostat* 1976;**2**(2): 99–110.
200. Bailey AG, Smedley CJA. The impact charging of polymer particles. *Adv Powder Technol* 1991;**2**(4):277–84.
201. Rowley G. Quantifying electrostatic interactions in pharmaceutical solid systems. *Int J Pharm* 2001;**227**(1–2):47–55.
202. Greason WD. Investigation of a test methodology for triboelectrification. *J Electrostat* 2000;**49**(3–4):245–56.
203. Oguchi T, Tamatani M. Contact electrification phenomena and powder surface treatments. *Wear* 1993;**168**(1–2):91–8.
204. Higashiyama Y, Ujiie Y, Asano K. Triboelectrification of plastic particles on a vibrating feeder laminated with a plastic film. *J Electrostat* 1997;**42**(1–2):63–8.
205. Luga A, Calin L, Neamtu V, Mihalcioiu A, Dascalescu L. Tribocharging of plastics granulates in a fluidized bed device. *J Electrostat* 2005;**63**(6–10):937–42.
206. Nguyen TTH, Hammond RB, Roberts KJ, Marziano I, Nichols G. Precision measurement of the growth rate and mechanism of ibuprofen {001} and {011} as a function of crystallization environment. *CrystEngComm* 2014;**16**:4568–86.
207. Ho R, Hinder SJ, Watts JF, Dilworth SE, Williams DR, Heng JYY. Determination of surface heterogeneity of d-mannitol by sessile drop contact angle and finite concentration inverse gas chromatography. *Int J Pharm* 2010;**387**(1–2):79–86.
208. Vasileiadis M, Pantelides CC, Adjiman CS. Prediction of the crystal structures of axitinib, a polymorphic pharmaceutical molecule. *Chem Eng Sci* 2015;**121**:60–76.
209. Shah UV, Olusanmi D, Narang AS, Hussain MA, Gamble JF, Tobyn MJ, Heng JYY. Effect of crystal habits on the surface energy and cohesion of crystalline powders. *Int J Pharm* 2014;**472**(1–2):140–7.
210. Kempkes M, Vetter T, Mazzotti M. Measurement of 3D particle size distributions by stereoscopic imaging. *Chem Eng Sci* 2010;**65**(4):1362–73.
211. Schorsch S, Hours J, Vetter T, Mazzotti M, Jones CN. An optimization-based approach to extract faceted crystal shapes from stereoscopic images. *Comput Chem Eng* 2015; **75**:171–83.
212. Winn D, Doherty MF. Modeling crystal shapes of organic materials grown from solution. *AIChE J* 2000;**46**(7):1348–67.
213. Lovette MA, Browning AR, Griffin WD, Sizemore JP, Snyder RC, Doherty MF. Crystal shape engineering. *Ind Eng Chem Res* 2008;**47**:9812–33.
214. Sizemore JP, Doherty MF. A new model for the effect of molecular imposters on the shape of faceted molecular crystals. *Cryst Growth Des* 2009;**9**(6):2637–45.

CHAPTER

9

Fundamental of Diffusion and Dissolution

Y. Chen[1], *J. Wang*[2] *and D.R. Flanagan*[3]

[1]Novast Laboratories, Ltd., Nantong, China [2]Actavis, Weston, FL, United States
[3]University of Iowa, Iowa City, IA, United States

9.1 FUNDAMENTAL OF DIFFUSION

9.1.1 Introduction

In nature, particularly in the gas and liquid phases, molecules are frequently colliding with conservation of momentum. Collisions cause changes in the direction of movement of molecules or small particles, resulting in their random motion. Collisions between molecules and the walls of a container generate pressure in a gas or hydrostatic pressure in a liquid. Random motion of a particle caused by random bombardment by solvent (or gas) molecules due to thermal motion is known as *Brownian motion*. *Diffusion* is the process by which atoms, molecules, or small particles are transported from a region of higher concentration to a region of lower concentration due to random motion. In the absence of convection, Brownian motion or diffusion is a primary transport mechanism for small particles ($<0.1\,\mu m$) or molecules when the transport distance is small (ie, $<$ a few millimeters). Over larger distances, diffusion is a slow process for material transfer, and other factors like convection are important to obtain significant transport of material in a reasonable time period (ie, <1 day). In dissolution testing, diffusion and convection are combined processes that generate dissolved drug in solution, in which convection is the dominating mechanism.

The diffusion process is driven by entropy change (ie, $\Delta S > 0$ for diffusion). The process is an example of the second law of thermodynamics; no enthalpy is required in the process (ie, no heat or work is required, but heat can affect the rate of diffusion). Diffusion is irreversible because it can occur in only one direction—from high to low concentrations.

The exact movement of any individual molecule (or small number of molecules) is completely random. The overall behavior (ie, macroscopic) of a large population of molecules ($>10^6$) is quite predictable. Diffusion is a process of mass transport that involves the movement of one molecular species through another (ie, a drug in a gel or solution). It occurs by random molecular motion from one point to another, and it can occurs in gas, liquid, or solid states. Diffusion in liquid, gel, or polymeric systems is of primary interest in the pharmaceutical sciences.

The subject of diffusion has been well studied, and a variety of models and their mathematical expressions related to pharmaceutical applications are available in the literature.[1–6] The first part of this chapter provides a review of the fundamental theory and the basic equations of diffusion to facilitate the further application of diffusion theory in pharmaceutical sciences.

9.1.2 Basic Equations of Diffusion

The two basic equations describing the diffusion process are Fick's first and second laws.

Fick's first law: The rate of transfer (dQ/dt) of a diffusing substance through the unit area (A) in one dimension (x) is proportional to the concentration gradient, $\partial C/\partial x$, as described in Eq. (9.1).

$$J = \frac{dQ}{Adt} = -D\frac{\partial C}{\partial x} \tag{9.1}$$

where:

J is the flux (rate of diffusion)
D is the diffusion coefficient
C is concentration
$\partial C/\partial x$ is the concentration gradient.

DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00009-1

The negative sign indicates that diffusion occurs from high concentration to low concentration.

Fick's second law: The rate of solute accumulation or concentration change, $\partial C/\partial t$, in an element volume, dV, is proportional to the flux gradient of the solute in the element, dJ/dV. In other words, the rate of change of number of molecules in the volume element equals the difference between the flux of molecules entering the element and the flux of molecules leaving the element.

The mathematical form of Fick's second law can be developed from the first law by the following approach: molecules entering and leaving a volume element by diffusion can occur in three dimensions. The contribution of diffusion in one of the dimensions, x, to the concentration change in the volume element is depicted in Fig. 9.1, where the surface area for diffusion remains constant along the diffusional path. By mass balance, the rate of change of molecules in the element by one dimension diffusion can be described by Eq. (9.2).

Thus, the concentration change in this volume element can be formulated as the difference in fluxes into and out of the element, as shown here:

$$\frac{\partial C}{\partial t}dV = (J_{\text{in}} - J_{\text{out}})dA \tag{9.2}$$

Utilizing the relationship of $dV = dAdx$ and replacing dx (in the flux expression) with ∂x, the rate of concentration change in one-dimensional diffusion can be written as

$$\frac{\partial C}{\partial t} = \frac{-\partial J}{\partial x} = D\frac{\partial^2 C}{\partial x^2} \tag{9.3}$$

Similar equations can be derived for diffusion in the other two dimensions, assuming that the diffusion coefficient in the volume element is constant. The rate of concentration changes can thus be written as the sum of contributions in three dimensions as

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} + D\frac{\partial^2 C}{\partial y^2} + D\frac{\partial^2 C}{\partial z^2} \tag{9.4}$$

Eqs. 9.2–9.4 are derived for the case where the surface area for diffusion remains constant. If the surface area for diffusion changes, such as the radial diffusion in a cylinder or in a sphere, the area changes must be taken into account. For diffusion in a sphere, Fick's second law is

$$\frac{\partial C}{\partial t} = D\left(\frac{\partial^2 C}{\partial r^2} + \frac{2}{r}\frac{\partial C}{\partial r}\right) \tag{9.5}$$

For radial diffusion in a cylinder, Fick's second law can be expressed as

$$\frac{\partial C}{\partial t} = D\left(\frac{\partial^2 C}{\partial r^2} + \frac{1}{r}\frac{\partial C}{\partial r}\right) \tag{9.6}$$

Diffusion can reach a steady state where concentration in a volume element does not change with time, but material still diffuses through the element. Using this criterion, the rate of change of the concentration gradient in the element is zero at steady state according to Fick's second law. For a one-dimensional system with constant surface area, the concentration gradient ($\partial C/\partial x$) at steady state is constant (the concentration profile is linear with distance along the diffusional path in the system).

If one views Fick's first law as an equation for the rate of mass transfer, then Fick's second law expresses the acceleration (or deceleration) of mass transfer.

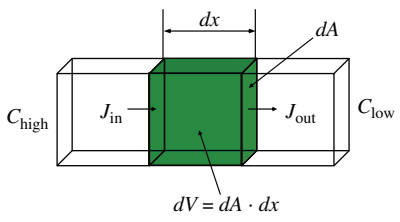


FIGURE 9.1 Volume element.

9.1.3 Solutions for Diffusion Equations

The most fundamental diffusion equation, Fick's second law, is a partial differential equation that must be solved for specific conditions before it can be used for practical problems. Mathematical solutions for actual systems can be obtained by taking the initial and boundary conditions of the real systems into consideration. Assuming that the diffusion coefficient is a constant, solutions for many systems with different initial and boundary conditions have been derived, either in the form of an error function (or a series of error functions) or of a Fourier series. These solutions can be found in the literature,[1–6] eliminating the need to derive them again. One can thus select the corresponding solution from the literature to fit the specific initial and boundary conditions for a particular system. These can be either infinite or finite systems, with or without perfect mixing (ie, sink conditions) for the boundary conditions. A number of these cases will be presented and discussed next.

9.1.3.1 Diffusion from a plane source into an infinite medium

If a finite amount of solute, M, is deposited on an infinitely thin plane in the middle of an infinite tube of medium (ie, water or gel), and the solute diffuses

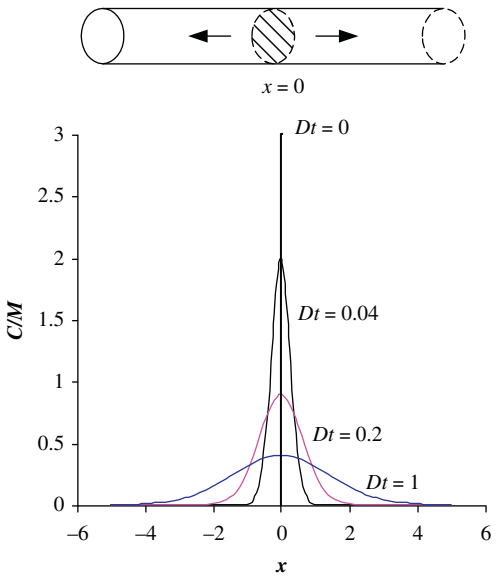


FIGURE 9.2 Plane-sourced, one-dimensional diffusion system and the relative concentration profiles as a function of distance and time.

axially toward each end of the tube (ie, one-dimensional diffusion) without mixing, as depicted in Fig. 9.2, Eq. (9.7) is the solution to Fick's second law for this one-dimensional system, assuming that the diffusion coefficient D is constant at constant temperature:

$$C(x,t) = \frac{M}{2\sqrt{\pi Dt}} \exp\left(\frac{-x^2}{4Dt}\right), \quad t > 0 \tag{9.7}$$

As indicated in Fig. 9.2, the concentration at $x = 0$ becomes smaller with time. At other points where $|x| > 0$, concentration increases with time when $\partial^2 C/\partial x^2 > 0$ and decreases when $\partial^2 C/\partial x^2 < 0$. The area under the concentration distribution curve is constant because the entire initial mass, M, is retained. However, as t approaches infinity, C becomes infinitely small over all distance values.

9.1.3.2 *Diffusion between two infinite regions in contact*

This system consists of two regions in contact with one of the regions (ie, $x > 0$) initially containing solute at concentration C_0, and the other region is free of solute (ie, $C = 0$). Solute diffuses from the initially loaded region into the empty region without mixing. This system can be considered an extended version from the plane source system described previously by treating the system as containing an infinite number of plane sources of thickness, $\Delta\alpha$, each with an initial concentration, C_0, as shown in Fig. 9.3.

The initial conditions are

$$t = 0, \quad C = 0 \quad \text{for } x < 0, \quad \text{and} \quad C = C_0 \quad \text{for } x > 0$$

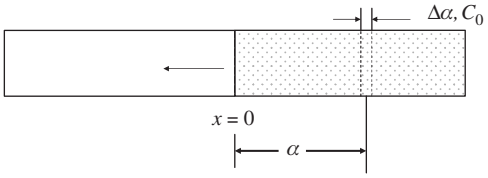


FIGURE 9.3 Infinitely extended region of uniform initial distribution.

Obviously, the amount of diffusing molecules in each segment is $C_0\Delta\alpha$, and diffusion from each infinitely thin segment to the point α can be expressed as follows:

$$C(x,t) = \frac{C_0\Delta\alpha}{2\sqrt{\pi Dt}} \exp\left(-\frac{(x-\alpha)^2}{4Dt}\right) \tag{9.8}$$

By superposition, diffusion from the entire region with initial distribution C_0 can be obtained by integrating over all the infinitely small elements to obtain:

$$C(x,t) = \int_0^\infty \frac{C_0}{2\sqrt{\pi Dt}} \exp\left(-\frac{(x-\alpha)^2}{4Dt}\right) d\alpha \tag{9.9}$$

Eq. (9.9) can be rewritten in the form of an error function:

$$C(x,t) = \frac{C_0}{2}\left[1 + erf\left(\frac{x}{2\sqrt{Dt}}\right)\right] \tag{9.10}$$

Eq. (9.10) shows that at the interface where $x = 0$, the concentration is constant at $C_0/2$.

The net amount of material transferred from the initially loaded region into the unloaded region (ie, across $x = 0$) over time can be calculated as follows:

$$Q_t = AC_0\sqrt{\frac{Dt}{\pi}} \tag{9.11}$$

where A is the cross-sectional surface area of the entire system. Note that Eqs. (9.10) and (9.11) are applicable for infinite regions in contact or a short time approximation of finite systems with the same initial conditions as the infinite systems. As a short time solution, it is assumed that the initially loaded region maintains C_0 at some distance from $x = 0$, and the initially unloaded region is still essentially free of solute at some distance from $x = 0$. The derivation of Eqs. (9.10) and (9.11), as well as the properties of the error function, were reviewed in the previous edition of this book.[7]

9.1.3.3 *Diffusion in semi-infinite systems*

If a semi-infinite polymeric gel or a large planar piece of biological tissue initially free of drugs is in

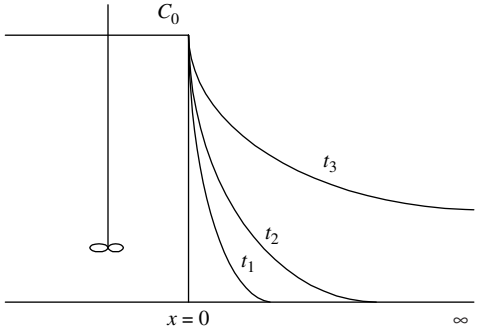


FIGURE 9.4 Diffusion in a semi-infinite system.

contact with a well-stirred saturated drug solution, as depicted in Fig. 9.4, the initial conditions are given by

$$C = 0, \quad x > 0, \quad t = 0$$
$$C = C_0, \quad x = 0, \quad t > 0$$

where C_0 is the drug solubility at the interface of the solution and the gel or tissue surface.

Under these conditions, the drug concentration profile that has diffused into the gel or tissue up to time t can be described by

$$C(x,t) = C_0\left\{1 - erf\left[\frac{x}{2\sqrt{Dt}}\right]\right\} \tag{9.12}$$

The amount of drug that has diffused into the gel over time can be described by

$$Q_t = 2AC_0\sqrt{\frac{Dt}{\pi}} \tag{9.13}$$

Comparing Eqs. (9.11) and (9.13), it can be seen that the rate of diffusion is doubled when half of the diffusion system is well mixed, and maintained at a constant concentration compared to the unstirred system, with all other conditions being the same.

In infinite and semi-finite systems, the concentration profiles are a function of diffusion distance and time; the thickness of the system does not affect the concentration profile until concentration begin to change at either boundary in the semi-infinite system.

9.1.3.4 Diffusion in finite planar systems

When a planar gel of thickness $2h$, initially loaded with a solute at $C = C_0$, is placed in a well-stirred medium maintained at zero concentration, drug release occurs with an initial condition of

$$C = C_0, \quad -h < x < h, \quad t = 0$$

and boundary conditions of

$$C = 0, \quad x = \pm h, \quad t > 0, \quad \text{and} \quad \frac{\partial C}{\partial x} = 0, \quad x = 0, \quad t > 0$$

For this finite system, the thickness of the gel affects drug release. The complete solution for the concentration in the gel as a function of time and distance is

$$C(x,t) = C_0\left\{1 - \sum_{n=0}^{\infty}(-1)^n erfc\left(\frac{(2n+1)h - x}{2\sqrt{Dt}}\right) - \sum_{n=0}^{\infty}(-1)^n erfc\left(\frac{(2n+1)h + x}{2\sqrt{Dt}}\right)\right\} \tag{9.14}$$

The amount of drug released from the gel over time is

$$Q_t = 2AC_0\sqrt{Dt}\left\{\frac{1}{\sqrt{\pi}} + 2\sum_{n=1}^{\infty}(-1)^n ierfc\left[\frac{nh}{\sqrt{Dt}}\right]\right\} \tag{9.15}$$

For a small period of time when $nh/\sqrt{Dt}$ is large, the *ierfc* terms become negligible, and Eq. (9.15) can be simplified to Eq. (9.13), which is useful for a drug release of not more than approximately 60%. This approximation can be understood by considering the fact that for short periods, the middle of the gel is still at C_0, so the system can be viewed as a semi-finite system. Therefore, Eq. (9.15) can be reduced to Eq. (9.13) for short time approximations, even though the systems differ in size.

Using the method of separation of variables,[7] the concentration in the finite planar systems can be expressed by a series of trigonometric distance and exponential time functions, known as *Fourier series*, as shown here:

$$C = \frac{4C_0}{\pi}\sum_{p=0}^{\infty}\frac{(-1)^p}{2p+1}\cos\frac{(2p+1)\pi x}{2h}\exp\left(-\frac{(2p+1)^2\pi^2 Dt}{4h^2}\right) \tag{9.16}$$

The total amount of drug released from both sides of the gel with a total surface area of $2A$ is described by

$$Q = \frac{8AC_0h}{\pi^2}\left\{\frac{\pi^2}{4} - \sum_{p=0}^{\infty}\frac{1}{(2p+1)^2}\exp\left(-\frac{(2p+1)^2\pi^2 Dt}{4h^2}\right)\right\} \tag{9.17}$$

Eq. (9.17) is useful for drug releases of more than approximately 60%, and can be considered a useful long-term solution. The Fourier series in this case is not a convenient short-term solution because it requires too many terms to be calculated in the series to obtain accurate concentration profiles, but at longer times, only a few terms (or possibly only the first term) are needed for accurate calculations.

9.1.3.5 Diffusion across a planar barrier

In membrane permeation studies, it is common to place a homogenous barrier membrane of thickness h

between a donor solution and a receiver solution. The solutions on both sides are well mixed such that the concentration is maintained constant at C_0 in the donor solution, and essentially at zero in the receiver solution (called *sink conditions*). Under these conditions, solute will diffuse from the donor solution through the barrier membrane and pass into the receiver solution. The initial condition for this system is

$$C = 0, \quad 0 < x < h, \quad t = 0$$

and the boundary conditions are

$$C = C_0, \quad x = 0, \quad \text{and} \quad C = 0, \quad x = h, \quad t > 0$$

Assuming a constant diffusion coefficient at constant temperature, the concentration profile of solute in the membrane can be found by separation of variables and proves to be

$$C = C_0 - \frac{xC_0}{h} - \frac{2C_0}{\pi}\sum_{n=1}^{\infty}\frac{1}{n}\sin\frac{n\pi x}{h}\exp\left(-\frac{n^2\pi^2 Dt}{h^2}\right) \tag{9.18}$$

At the early stage of diffusion when time is short, the concentration gradient in the membrane is not yet steady, and it changes with respect to time and distance. During this stage, sorption of solute by the membrane is the dominant process because the rate of solute entering the membrane is much greater than the rate leaving the membrane on the receiver side. At longer times, such as $t \to \infty$, the sorption rate by the membrane reaches a constant, and the rate of solute entering the membrane is the same as that leaving the membrane. At this point, the solute concentration profile in the membrane becomes linear with distance and does not change with time. This stage is commonly referred to as *steady state*.

Regardless of the time, the general equation for the amount of solute passing through unit surface area of membrane up to time t can be calculated by deriving the expression for the concentration gradient (dC/dx), calculating the flux (J) at $x = h$, and then integrating the flux at this point over time. The resulting equation is:

$$Q = \frac{DC_0 t}{h} - \frac{hC_0}{6} - \frac{2hC_0}{\pi^2}\sum_{n=1}^{\infty}\frac{(-1)^2}{n^2}\exp\left(-\frac{n^2\pi^2 Dt}{h^2}\right) \tag{9.19}$$

As $t \to \infty$ as steady state is reached, the exponential term in Eq. (9.19) becomes negligible and the rate of solute diffusing through the membrane becomes constant with time, as described by the first two terms of Eq. (9.19), shown here:

$$Q = \frac{DC_0}{h}\left(t - \frac{h^2}{6D}\right) \tag{9.20}$$

By extrapolating the linear profile back to $Q = 0$, an intercept on the time axis is obtained:

$$t_L = \frac{h^2}{6D} \tag{9.21}$$

t_L is commonly referred to as the *lag time*. Steady state is reached asymptotically and is within about 99% of exact steady state at $t = 2.7t_L$, which can be used as a guideline for membrane permeation studies to ensure that steady state is reached, if such a state is desired.

Eqs. (9.18)–(9.21) in fact apply to a unique case in which the membrane is initially free of solute. In general, the membrane may initially contain a solute at a concentration of C_i such that the initial conditions are

$$C = C_i, \quad 0 < x < h, \quad t = 0$$
$$C = C_0, \quad x = 0; \quad C = 0, \quad x = h, \quad t \geq 0$$

Crank[2] showed that the general equation for the amount of solute that has diffused through a unit area of membrane is

$$Q = \frac{DC_0}{h}\left(t + \frac{C_i h^2}{2DC_0} - \frac{h^2}{6D}\right) - \frac{2hC_0}{\pi^2}\sum_{n=1}^{n=\infty}\frac{(-1)^n}{n^2}\exp\left(-\frac{n^2\pi^2 Dt}{h^2}\right) - \frac{4C_i h}{\pi^2}\sum_{m=0}^{m=\infty}\frac{1}{(2m+1)^2}\exp\left(\frac{-(2m+1)^2\pi^2 Dt}{h^2}\right) \tag{9.22}$$

If $C_i = C_0$, as $t \to \infty$, Eq. (9.22) approaches the linear function:

$$Q = \frac{DC_0}{h}\left(t + \frac{h^2}{3D}\right) \tag{9.23}$$

A burst time t_B (actually, this is also a lag time) can be obtained by extrapolating the linear release profile to $Q = 0$:

$$t_B = \frac{-h^2}{3D} \tag{9.24}$$

9.1.3.6 Diffusion in a sphere

Diffusion in a sphere is a three-dimensional process. The surface area for diffusion changes as mass diffuses radially. Taking this geometry into consideration, the differential equation for Fick's second law of diffusion in a sphere is

$$\frac{\partial C}{\partial t} = D\left(\frac{\partial^2 C}{\partial r^2} + \frac{2}{r}\frac{\partial C}{\partial r}\right) \tag{9.25}$$

Setting $u = Cr$, it can be shown that

$$\frac{\partial u}{\partial t} = r\frac{\partial C}{\partial t}(\partial r/\partial t = 0),$$
$$\frac{\partial^2 u}{\partial r^2} = 2\frac{\partial C}{\partial r} + r\frac{\partial^2 C}{\partial r^2},$$

and Eq. (9.25) can be written as

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial r^2} \tag{9.26}$$

This is exactly the same form as Eq. (9.3) and can be solved in the same fashion as for one-dimensional planar diffusion.

For a homogeneous permeable sphere of radius R, initially free of solute and placed in a well-stirred solution such that $C = C_0$ at $r = R$, the concentration profile in the sphere is given by

$$C = C_0\left\{1 - \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{R}{r}\frac{(-1)^{n+1}}{n}\sin\frac{n\pi r}{R}\exp\left(-\frac{n^2\pi^2 Dt}{R^2}\right)\right\} \tag{9.27}$$

The total amount of solute that has entered the sphere up to time t is

$$Q = \frac{8R^3C_0}{\pi}\left\{\frac{\pi^2}{6} - \sum_{n=1}^{\infty}\frac{1}{n^2}\exp\left(-\frac{n^2\pi^2 Dt}{R^2}\right)\right\} \tag{9.28}$$

9.1.3.7 Diffusion in a cylinder

Assuming that axial diffusion through the two ends of a cylinder is negligible, the radial diffusion in a cylinder is a two-dimensional diffusion system. The partial differential equation for this case has been given by Eq. (9.6). Similar to the one-dimensional diffusion problems, solutions for radial diffusion in a cylinder can be derived using an error function or by separation of variables, which results in Bessel function solutions. These solutions may appear more complicated than those derived for the one-dimensional diffusion systems, but there are expressions available for numerous cylindrically symmetrical systems. Interested readers are referred to the research literature for detailed mathematical treatments for general solutions. A simple case discussed here is the steady-state radial diffusion through the wall of a hollow cylinder.

For outward diffusion through the wall of a hollow cylinder with inner radius a and outer radius b, and with the concentration profile maintained constant at C_0 at $r = a$ and $C = 0$ at $r = b$, the steady-state expression is

$$D\left\{\frac{\partial^2 c}{\partial r^2} + \left(\frac{1}{r}\right)\frac{\partial c}{\partial r}\right\} = 0 \tag{9.29}$$

The general solution is

$$C = A + B\ln r \tag{9.30}$$

From the boundary conditions, it can be determined that

$$B = \frac{C_0}{\ln(a/b)}, \text{ and } A = -\frac{C_0\ln b}{\ln(a/b)}$$

Therefore, the steady-state concentration profile in the wall is

$$C = \frac{C_0\ln(r/b)}{\ln(a/b)} \tag{9.31}$$

The quantity of drug released (Q) through $r = b$ into a perfect sink as a function of time is given by

$$Q = \frac{ADC_0 t}{\ln(a/b)} \tag{9.32}$$

Eq. (9.32) shows that, at steady state, the rate of diffusion through the wall of a hollow cylinder is constant, as expected, but the concentration profile in the wall as given by Eq. (9.31) is nonlinear.

9.1.3.8 Diffusion combined with other processes

Diffusion and adsorption

Adsorption of diffusing molecules in some systems, such as in a reinforced polymer matrix or onto solids in suspensions, can occur during the diffusion process. For example, silicon dioxide or other fillers in reinforced polymeric membranes can adsorb the diffusant during membrane permeation. Adsorption leads to the immobilization of some of the diffusing substance, hence reducing the driving force and diffusion rate until adsorption reaches equilibrium with the free solute concentration at each position in the membrane. Typically, the amount of substance immobilized by the adsorbent is determined by the adsorption isotherm, which is a function of the free concentration of the diffusant. To account for the adsorption, Fick's second law for one-dimensional diffusion is modified to give

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} - \frac{V_f}{\varepsilon}\frac{\partial A_f}{\partial t} \tag{9.33}$$

where: A_f represents the sorption isotherm of the adsorbent, V_f is the volume fraction of adsorbent in the membrane, and ε is the porosity of the continuous polymeric phase.

In this case, the lag time for diffusion has been shown[8] to be

$$t_L = \frac{\tau^2 h^2}{6D} + \frac{1}{DC_0}\frac{V_f}{\varepsilon}\int_0^{\tau h} xA_f d_x \tag{9.34}$$

where: h is the thickness of the membrane and τ is the tortuosity due to the presence of the impermeable filler.

Eq. (9.34) shows that lag time for membrane diffusion can be increased significantly due to increasing tortuosity and adsorption by the filler. The effects of filler on lag time have also been verified by experimental results. Chen studied the diffusion of 2-hydroxypyridine through polydimethylsiloxane (PDMS) membranes with different levels of silica filler. It was shown that lag-time increased by about 10-fold by adding 23% silica filler into the PDMS membrane with thickness similar to the unfilled one.[9] The lag time calculated agreed well with the predicted value from Eq. (9.34). The filler effect on steady-state flux, on the other hand, is substantially less than the effect on lag time. After adsorption reaches equilibrium at each point in the membrane, the steady-state flux through the membrane containing 23% of the filler is about 50% of the unfilled one (Fig. 9.5).

Diffusion and chemical reaction

If a slow chemical reaction occurs simultaneously with diffusion, the reaction reduces the free diffusant concentration. Similar to the adsorption problem, the expression of Fick's second law is reduced by the reaction rate, R, as described by

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} - R \tag{9.35}$$

If the rate of reaction is simple (ie, first-order), and is directly proportional to the rate of concentration change as a function of time ($R = k\partial C/\partial t$), Eq. (9.35) can be rewritten as

$$\frac{\partial C}{\partial t} = \frac{D}{1+k}\frac{\partial^2 C}{\partial x^2} = D'\frac{\partial^2 C}{\partial x^2} \tag{9.36}$$

where D' is the reduced diffusion coefficient due to reaction. It can be seen from Eq. (9.36) that the mathematical solution for diffusion coupled with a simple first-order reaction is the same as a simple diffusion process without reaction, except that the diffusion coefficient is reduced. It should be noted that if $k \gg D$, diffusion may not be detectable.

For reaction rates other than this simple case, more complex solutions can be developed based on the applicable rate law and the initial and boundary conditions for the diffusion-reaction process. Readers are referred to the research literature[1–6] for details on more complicated diffusion-reaction problems.

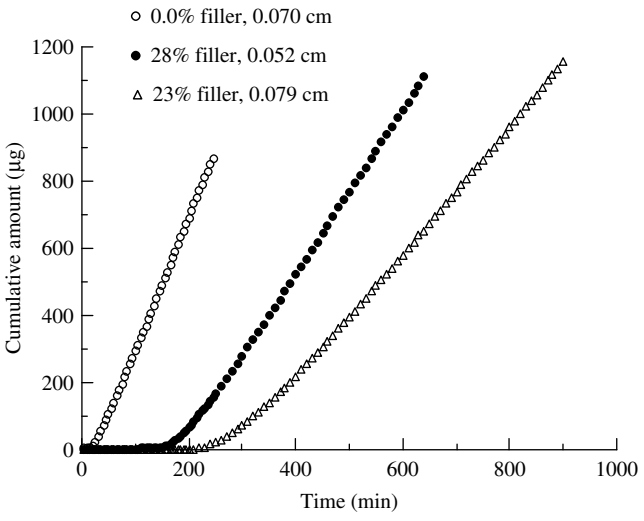


FIGURE 9.5 Flux of 2-hydroxypyridine through polydimethylsiloxane membranes with different levels of silica filler.

9.1.4 The Diffusion Coefficient and Its Determination

From Fick's first law, the diffusion coefficient D can be defined as the rate of transfer of the diffusing substance across unit area and unitary concentration gradient of the diffusing substance. From this definition, it can be shown that the diffusion coefficient has units of length^2/time, which is commonly expressed in cm^2/s. The diffusion coefficient is a property of a solute in a specific medium or matrix at a constant temperature and pressure. It is dependent on temperature, viscosity, and the size of solute, as described by the Einstein equation

$$D = \frac{k_B T}{6\pi\eta r} \tag{9.37}$$

where:

k_B is the Boltzmann's constant
T is absolute temperature
η is the medium viscosity
r is the solute radius.

If all the values of the parameters on the right side of Eq. (9.37) are known, the diffusion coefficient can be calculated.

For a given solute in a given medium, the dependency of diffusion coefficient on temperature is described by

$$D = D_0 \exp\left(-\frac{E_a}{RT}\right) \tag{9.38}$$

where D_0 is the hypothetical diffusion coefficient at infinite temperature and E_a is the activation energy of diffusion.

Assuming that the diffusion coefficient is not affected by solute concentration, the diffusion

coefficient can be determined from steady-state fluxes or lag times from a membrane permeation study, or from release data of other well-defined systems at constant temperature. These methods are discussed next.

9.1.4.1 Steady-state flux method

As discussed earlier, for membrane permeation with constant surface concentration, diffusion reaches steady state when the solute concentration profile in the membrane becomes linear with respect to distance. At steady state, the flux is proportional to the diffusion coefficient and concentration gradient as shown here:

$$J_{ss} = Dk\frac{\Delta C}{h} \tag{9.39}$$

where:

J_{ss} is the steady-state flux
k is the partition coefficient
ΔC is the concentration difference between the donor and the receiver solutions.

If membrane permeation is conducted in aqueous solution, the membrane may be covered by stagnant aqueous layers at the membrane-solution interfaces, either due to a significant partition rate limitation between the membrane and the aqueous solutions or the poor hydrodynamics of mixing at the interfaces. Assuming that the thicknesses of the stagnant layers on both sides of the membrane are the same, a general flux equation is given[10] by

$$J_{ss} = \frac{\Delta C}{h_m/D_m k + 2h_a/D_a} = \frac{\Delta C}{(1/P_m) + (2/P_a)} \tag{9.40}$$

where the subscripts m and a indicate membrane and aqueous solutions, respectively; and P is the permeability coefficient.

Eq. (9.40) shows that flux could be determined for the membrane with the stagnant diffusion layers, and depending on the properties of the membrane (ie, thickness and partition coefficient), the aqueous diffusion layers may or may not contribute significantly to the overall flux. If $h_m/D_m k \gg 2h_a/D_a$, diffusion is controlled by the membrane. If $h_m/D_m k \ll 2h_a/D_a$, diffusion is controlled by the stagnant diffusion layers. It has been reported that under some common study conditions, the thickness of the stagnant aqueous diffusion layers are about 100 μm. However, the thickness of these stagnant layers decreases as the agitation rate increases. If a membrane has a sufficient thickness and the study is conducted with a high stirring rate, the boundary layer effects may be negligible, although this assumption must be verified by experimentation. Once it is proved that the boundary layer effects are negligible, the diffusion coefficient in a homogeneous polymer membrane can be calculated from the steady-state flux data, as shown here:

$$D_m = \frac{J_{ss}h}{k\Delta C} \tag{9.41}$$

If the effect of the boundary layer is not negligible, additional experimentation must be conducted in order to obtain a reliable membrane diffusion coefficient. By determination of the total apparent permeability at constant temperature and constant stirring rate for membranes of different thicknesses,[10] the intrinsic membrane permeability can be calculated from the intercept of a graph of the reciprocal of apparent permeability coefficient P_{total} against the reciprocal of membrane thickness, as described by

$$\frac{1}{P_{total}} = \frac{1}{P_m} + \frac{2h_a}{D_a}\frac{1}{h_m} \tag{9.42}$$

In order to obtain accurate diffusion coefficients, it is also important to carry out the membrane permeation experiment for a sufficiently long time to ensure that a steady state is reached. As a general guide, flux data used for calculation should be determined to be greater than 2.7 × lag time.

If the membrane is heterogeneous with impermeable inert domains, the measured value from steady-state flux data is an apparent diffusion coefficient of the composite material. The volume fraction (V) and the tortuosity (τ) of the continuous polymer phase must be accounted for in order to calculate an accurate diffusion coefficient of the continuous membrane phase according to

$$D_{eff} = D\frac{V}{\tau} \tag{9.43}$$

9.1.4.2 Lag time method

As discussed previously, when a membrane initially free of solute is used to conduct a permeation study with the concentration in the donor solution kept constant, and the receiver solution is kept at essentially zero concentration, the diffusion coefficient can be calculated from the lag time by

$$D = \frac{h^2}{6t_L} \tag{9.44}$$

If the membrane contains inert fillers, lag time is increased by the tortuosity, which must be considered for the calculation of diffusion coefficient as follows:

$$D = \frac{\tau^2 h^2}{6t_L} \tag{9.45}$$

Care should be taken when a thin membrane is used to conduct diffusion research in aqueous solution. As

discussed previously, diffusion may be controlled by the aqueous boundary diffusion layers for some systems. In the extreme case when $h_m/D_m k \ll 2h_a/D_a$, the lag time is determined by the boundary layers as

$$t_L = \frac{2h_a^2}{3D_a} \tag{9.46}$$

Obviously, the diffusion coefficient of the membrane could not be calculated from the lag time when the process is controlled by the stagnant layers.

If the membrane is initially equilibrated with the donor solution prior to the initiation of permeation study, there will be a negative burst time t_B, which is also lag time, and the diffusion coefficient can be calculated by

$$D = \frac{-h^2}{3t_B} \tag{9.47}$$

When the membrane contains adsorptive fillers or some other domains capable of binding the diffusing substance, the lag time is the combined result of diffusion in the polymer and adsorption or binding in the system; hence, Eqs. (9.44)–(9.46) are no longer valid for calculation of the diffusion coefficient. This is particularly important for biological membranes because in many cases, protein in biological membrane can bind many kinds of organic compounds. For adsorptive and binding systems, the effects of adsorption or binding on lag time (and hence the diffusion coefficient) must be accounted for by using Eq. (9.34) or modified versions of it.

If the adsorption isotherm of the filler is a simple linear equation as a function of the steady-state concentration distribution (C_s) in the membrane as follows:

$$A_f = bC_s = bC_0\left(1 - \frac{x}{h}\right) \tag{9.48}$$

then Eq. (9.34) can be integrated. The result of the corrected lag-time relationship with tortuosity and adsorption due to the presence of filler is given by

$$t_L = \frac{\tau^2 h^2}{6D}\left(1 + b\frac{V_f}{\varepsilon}\right) \tag{9.49}$$

Thus, the diffusion coefficient can be expressed as

$$D = \frac{\tau^2 h^2}{6t_L}\left(1 + b\frac{V_f}{\varepsilon}\right) \tag{9.50}$$

9.1.4.3 *Sorption and desorption methods*

Data from either absorption by a planar polymer matrix or release/desorption from the matrix have frequently been used to determine the diffusion coefficient in the polymer. The exponential series given by Eq. (9.17) can be used to determine the diffusion coefficients by curve fitting.

The short time approximation for Eq. (9.15), as in Eq. (9.13) using the initial sorption or desorption data, is more convenient to use for the determination of the diffusion coefficient than from using Eq. (9.17). From the slope of the Q_t versus $\sqrt{t}$ plot for the early release or sorption data, the diffusion coefficient can be calculated as follows:

$$D = \frac{\pi k^2}{4C_0^2} \tag{9.51}$$

where:

k is the slope of the Q_t versus $\sqrt{t}$ plot
C_0 is the initial concentration in the matrix in the case of release, or the surface concentration in the case of sorption.

For a heterogeneous system with impermeable domains, Eq. (9.51) must be corrected for porosity and tortuosity as follows:

$$D = \frac{\tau\pi k^2}{4\varepsilon^2 C_0^2} \tag{9.52}$$

Theoretically, different methods should yield the same value of diffusion coefficient for the same solute in the same polymer. However, if the polymer contains adsorptive domains, the results may not agree with each other, due to the difference in the effective concentration of the solute available for diffusion in the polymer phase at the time of the experiment. This is because adsorption or binding may be effective during the initial sorption study, but not during the early release stage. Therefore, it is important to understand and characterize the system before valid conclusions can be drawn.

9.1.5 Pharmaceutical Application of Diffusion Theory

Much study has been conducted on the application of diffusion theory in different fields of pharmaceutical science, especially in the study of controlled release drug delivery systems, dissolution of particles and dosage forms, and passive drug absorption processes. Vast information is available in the literature for numerous specific cases. A review of detailed applications is beyond the scope of this chapter. However, these authors believe that it is important to acknowledge several seminal applications of diffusion theory in pharmaceutical science for the study of drug delivery and dissolution of particles, as these applications have led to the development of models that are

playing critical roles in advancing the pharmaceutical sciences. For drug delivery, the application of the fundamental equations of Fick's first and second laws enabled the development of different equations to describe the drug release profiles for systems of different geometries and conditions, such as the Higuchi equations[11,12] and the Langer equation.[13] Similarly, several models were developed for studying the dissolution rates of particles with different particle sizes, including the diffusion layer model by Noyes and Whitney[14] and Nernst,[15] the cube-root model for describing the dissolution profiles of large particles by Hixson and Crowell,[16] the two-thirds-root equation for dissolution of fine particles by Higuchi and Hiestand,[17] and the unified general model developed by Wang and Flanagan[18] for describing the dissolution profiles of small and large spherical particles. The application of diffusion theory for the development of different equations for particle dissolution will be discussed in more detail in the following parts of this chapter.

9.2 FUNDAMENTALS OF DISSOLUTION

9.2.1 Introduction

Orally administered solid dosage forms generally undergo disintegration, dissolution, and absorption to enter the bloodstream and reach the site of action. For drugs with low solubility, dissolution is often the rate-limiting step that directly affects the rate and extent of drug absorption. Various dissolution tests have been developed and have become tools for predicting in vivo performance and monitoring product quality. The basic theoretical aspects of solid dissolution are covered in this chapter, and this discussion is focused on the dissolution of small drug molecules in aqueous media.

9.2.2 Mechanism and theories of solid dissolution

A prerequisite for drug molecules to be transported into the bulk solution is that they detach from the solid surface and form solvated molecules. This process is called *solvation*, which is a pseudochemical reaction. Solvated molecules then need to transfer from the solid–liquid interface to the bulk solution, which is a mass transport process. Both the solvation and the mass transport process can control the dissolution rates. Diffusion usually plays a key role in the transport process, while convection also is important in cases of agitated media.

9.2.2.1 *Thermodynamic considerations*

Whether a solid is crystalline or amorphous, neighboring molecules are closely associated with each other through intermolecular forces. Dissolution involves the breaking of these existing intermolecular interactions and the formation of new interactions with the solvent (usually water). The solid dissolution process generally involves two sequential steps. The first step is the interaction between solid and solvent molecules to form solvated molecules (solvation), which takes place at the solid–liquid interface. The second step is the mass transport of solvated molecules from the solid–liquid interface to the bulk solution. These two steps govern the rate and extent of solid dissolution. Solubility, a basic property for solids, controls the first step. Transport, on the other hand, usually controls the second step, which is generally slower than the first step. Overall, dissolution is governed by solubility and transport processes.

The first step of dissolution can be viewed as a pseudochemical reaction that is governed by the same principles as regular chemical reactions. The solvation process is reversible, and solubility is reached when the reaction reaches equilibrium. This process can be described by its Gibbs free energy (ΔG):

$$\text{Drug} + \text{Solvent} \leftrightarrow \text{Solvated Drug} \tag{9.53}$$

$$\Delta G = \Delta H - T\Delta S \tag{9.54}$$

This reaction involves both the breakage of existing intermolecular interactions (drug-drug, solvent-solvent) and the formation of new intermolecular interactions (drug-solvent). The net entropy (ΔS) change for this reaction is generally positive (dissolution causes an increase in disorder), which is favorable for dissolution. Therefore, the extent of dissolution is often determined by the net enthalpy change (ΔH). If the net enthalpy change is negative or close to zero, the reaction will continue until all the available solids are dissolved. The empirical rule, "like dissolves like," therefore, has a thermodynamic basis. On the other hand, if the net enthalpy change (ΔH) is positive, the reaction will progress until equilibrium is reached ($\Delta G = 0$). A saturated solution is obtained.

The reaction rate V is described kinetically by

$$\begin{aligned} V &= k_1[\text{Drug}][\text{Solvent}] \\ k_1 &= A\mathrm{e}^{-E_a/RT} \end{aligned} \tag{9.55}$$

The overall rate of dissolution is controlled by the slower step of the two-step dissolution process. Thermodynamically, the activation energy E_a for the formation of a solvated drug should be less than the energy to break water-water or drug-drug intermolecular interactions. Since hydrogen bonding is among the strongest intermolecular interactions, the activation

energy for reaction (Eq. 9.53) in many cases should be less than the energy of the water-water hydrogen bond, which is about 19 kJ/mol. In liquid water, hydrogen bonds are constantly being broken and reformed because thermal energy is sufficient to cause cleavage of these interactions. Therefore, at room temperature, the rate of solvation is generally so fast that equilibrium is also instantaneous. Compared to the rate of solvation, mass transport from the solid–liquid interface to bulk solution involves molecules moving a great distance compared to molecular dimensions, which is usually much slower and becomes the rate-limiting step.

9.2.2.2 *Dissolution by pure diffusion*

For solid dissolution processes, convection generally accompanies diffusion in transporting dissolved molecules. Dissolution through pure diffusion without any contribution from some form of convection is rare. Even without mixing, density gradients caused by concentration differences or thermal effects due to the heat of solvation can lead to natural convection. However, we can consider a hypothetical situation in which dissolution is purely diffusional.

Assume that a small drug particle is placed at the bottom of a large beaker full of water. The water is kept unstirred so that there is no convection. The drug particle may dissolve by diffusion, which is equivalent to having a diffusion layer of growing thickness. In this case, dissolution becomes a pure diffusional phenomenon.

When a particle dissolves by pure diffusion, the concentration at every point away from the solid–liquid interface rises, but the concentration gradient at any distance from the particle decreases with time. The overall dissolution rate of the particle becomes slower. Depending on the solubility and solvent volume, there can be two possible results. If the amount needed to saturate the solvent is equal to or more than the particle weight, the particle will eventually completely dissolve; if the amount needed to saturate the solvent is less than the particle weight, a saturated solution will eventually form, which coexists with the smaller particle.

No matter what the final result is, we will not be able to see the formation of a pseudo-steady state in this dissolution process. A pseudo-steady state can be reached if during some finite time in the dissolution process, the concentration gradient becomes relatively constant with time. The existence of a pseudo-steady state would make the dissolution rate expression simpler mathematically. However, in this case, the concentration gradient may be constantly changing, and we may not be able to apply the pseudo-steady state treatment. Later in this chapter, we will see cases where a pseudo-steady state is applicable.

9.2.2.3 *Diffusion layer model*

Noyes and Whitney first proposed the basic transport-controlled model for solid dissolution.[14] They suggested that when surface area is constant, the dissolution rate is proportional to the difference between solubility and the bulk solution concentration:

$$\frac{dQ}{dt} = k(C_s - C_b) \tag{9.56}$$

where: dQ/dt is the dissolution rate, k is a constant (mass transfer coefficient), C_S is solubility, C_b is bulk solution concentration.

Nernst[15] and Brunner[19] then suggested that rapid equilibrium (ie, saturation) is achieved at the solid–liquid interface during dissolution, and then diffusion occurs across a thin layer of stagnant solution, called the *diffusion layer*, into the bulk solution. Diffusion across this diffusion layer is rate-controlling in most cases, which effectively converts the heterogeneous dissolution process into a homogeneous process of liquid-phase diffusion. This model, as shown in Fig. 9.6, is also called the *film model*. Dissolution rate can be expressed as

$$\frac{dQ}{dt} = DA\frac{(C_s - C_b)}{h} \tag{9.57}$$

where:

D is diffusion coefficient,
A is solid surface area,
h is diffusion layer thickness.

Nernst's and Brunner's concept of a diffusion layer being a stagnant or unstirred layer of liquid adhering to the solid surface is hydrodynamically unrealistic but allows the complex dissolution processes to be analyzed in a tractable fashion. The diffusion layer model is widely used due to its simplicity and applicability to

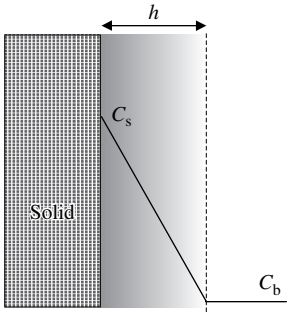


FIGURE 9.6 Diffusion layer model for a planar surface dissolution. (C_S is solubility, C_b is bulk solution concentration, h is diffusion layer thickness.)

a wide range of dissolution phenomena. However, the notion of a stagnant liquid layer near the solid surface is not supported by fluid dynamics. Liquid motion was observed at distances from solid surfaces that are much closer than with a typical diffusion layer thickness. Therefore, the diffusion layer is actually a boundary layer with a concentration gradient, as well as a velocity gradient. Another limitation of the diffusion layer model is that it does not give a means to calculate a diffusion layer thickness independently based on hydrodynamics. Thus, it does not allow computation of the diffusional flux, a priori. Diffusion layer thicknesses can be obtained only by fitting equations to experimental data.

9.2.2.4 Convective-diffusion model

Due to the limitations of the diffusion layer model, further examination of the dissolution process is necessary in order to develop an improved model. There are two mechanisms for the transport of solute molecules from a solid in a moving liquid. First, there is molecular diffusion as a result of a concentration gradient; second, solute molecules are carried along by the moving solvent (ie, convection). Dissolution is usually a combination of these two mechanisms, which is called *convective diffusion*. In convective-diffusion models, the contribution of both convection and molecular diffusion are taken into account. The general equation for convective diffusion in three dimensions is given here:

$$\frac{\partial C}{\partial t} = D\left(\frac{\partial^2 C}{\partial x^2} + \frac{\partial^2 C}{\partial y^2} + \frac{\partial^2 C}{\partial z^2}\right) - v_x\frac{\partial C}{\partial x} - v_y\frac{\partial C}{\partial y} - v_z\frac{\partial C}{\partial z} \tag{9.58}$$

where v_x, v_y, and v_z are liquid velocities in the x-, y-, and z-directions in Cartesian coordinates.

This equation is an extension of Fick's second law of diffusion with the addition of convection terms. It is a second-order partial differential equation with variable coefficients. To determine the net concentration change at a particular point at any time, we also need to know the liquid flow pattern and the boundary and initial conditions. This equation can be solved exactly for only a very limited number of special cases. Some of these cases will be discussed later in this chapter.

In the convective-diffusion model, solvent stream movement depends on its distance from the solid surface. At distances far from the solid surface, convection dominates over diffusion for transport of solute molecules. At distances very close to the solid surface, the solvent stream movement is slow, while the concentration gradient is large. In such regions, the contribution of diffusion to the solute molecule movement is comparable to or larger than convection. This region is also called the *diffusion layer*, but it is different from Nernst's concept. Levich pointed out the following differences between Nernst's layer and the convective diffusion-based diffusion layer[20]:

1. Solvent in Nernst's layer is considered stagnant, while it has a velocity gradient in the convective-diffusion model.
2. In Nernst's model, mass transport in the direction parallel to solid surface is disregarded. In convective-diffusion theory, convection and diffusion both across and along the diffusion layer are taken into account.
3. Nernst's layer is considered to have a constant thickness everywhere, while the thickness of the diffusion layer based on convective diffusion theory varies substantially along the solid surface depending on position and solvent motion. In fact, the diffusion layer has no clear-cut boundary; it is simply a region where the concentration gradient is at a maximum.

Unlike Nernst's diffusion layer model, the convective-diffusion model is more hydrodynamically realistic, which does allow us to predict solid dissolution rates. However, it is usually complicated to apply convective-diffusion to derive accurate expressions for solid dissolution because of the need to describe the fluid dynamics accurately.

9.2.3 Planar surface dissolution

Modeling planar surface dissolution is easier than modeling particulate dissolution because the mathematics is simpler. The simplest expression describing planar surface dissolution is the empirical Noyes-Whitney equation. The diffusion layer model is almost as simple. Measuring dissolution rate from a well-defined surface has been a popular way to characterize pharmaceutical solids. Determination of intrinsic dissolution rate is one example. Although the Noyes-Whitney equation is capable of describing such a process, planar surface dissolution can be more accurately described with convective-diffusion models based on different hydrodynamic conditions.

9.2.3.1 Convective-diffusion model for a rotating disk

When other parameters, such as agitation intensity and surface area, are fixed, dissolution rate can be viewed as a property of the solid. This loosely defined property is called the *intrinsic dissolution rate*. Due to the close relationship between the intrinsic dissolution rate and solubility, measuring the intrinsic dissolution rate becomes an alternative method for solubility estimation when equilibrium solubility cannot be easily

obtained experimentally. For example, measurement of intrinsic dissolution rates is a powerful tool for the evaluation of solubility differences between polymorphs or solvates. The equilibrium solubility of an anhydrous crystal form may not be determined if it converts to a hydrate when in contact with water. The solubility difference of an anhydrous versus hydrate crystal form can be estimated by studying the difference in initial intrinsic dissolution rates as the conversion occurs.

Since dissolution rate is always affected by agitation intensity, surface area, and container configuration, determination of intrinsic dissolution rates requires a method with good reproducibility. The rotating disk method, which is schematically demonstrated in Fig. 9.7, is often the method of choice. The disk is compressed from pure material and has a planar surface. The compressed disk is put into the holder so that only one flat surface is in contact with the solvent. The disk is then rotated at either a constant or varying speed, and the dissolution rate is measured. When a diffusion layer model is applied, a diffusion layer thickness can be easily determined from the dissolution rate data.

The hydrodynamics for a rotating disk system presents a good case of applying a convective-diffusion model. According to Levich, using a convective-diffusion model, the solvent flow pattern can be depicted as shown in Fig. 9.8.[21,22] Far from the disk, the fluid moves axially toward the disk. In a thin layer near the disk surface, the liquid acquires a rotating motion and moves radially away from the center of the disk. This system is one of the few cases that can be solved exactly with convective-diffusion theory. The dissolution rate (dQ/dt) is given by

$$\frac{dQ}{dt} = 1.95D^{2/3}C_s u^{-1/6}\omega^{1/2}R^2 \tag{9.59}$$

where:

D is the diffusion coefficient,
C_S is the solubility,
u is the kinematic viscosity,
ω is the rotational speed,
R is the disk radius.

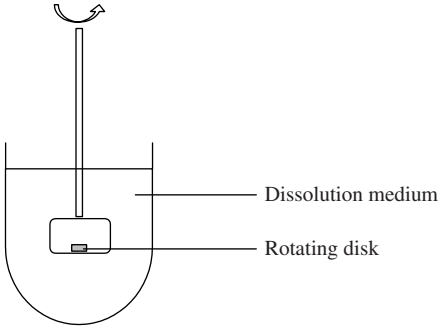


FIGURE 9.7 Rotating disk apparatus for measuring intrinsic dissolution rate.

The convective-diffusion model for dissolution of a rotating disk also provides a means to estimate the diffusion layer thickness as shown here:

$$h = 1.61D^{1/3}u^{1/6}\omega^{-1/2} \tag{9.60}$$

An interesting study using the rotating disk is by McNamara and Amidon, who applied convective-diffusion models to reactive dissolution.[23,24]

9.2.3.2 *Convective-diffusion model for flow past a planar surface*

In addition to dissolution from a rotating disk, there is the case of planar surface dissolution that can be modeled with relative ease. When the solvent flows parallel to a flat surface of a solid, there is a velocity gradient, as shown in Fig. 9.9. Both molecular diffusion normal to the surface and the forced convection

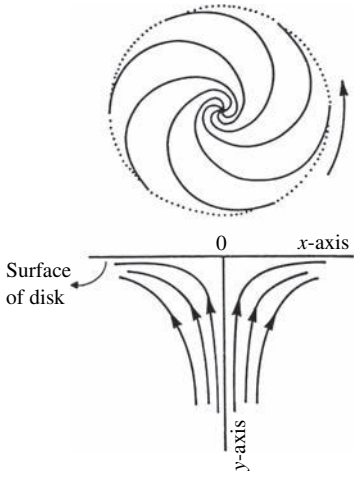


FIGURE 9.8 Liquid flow pattern around a rotating disk apparatus.

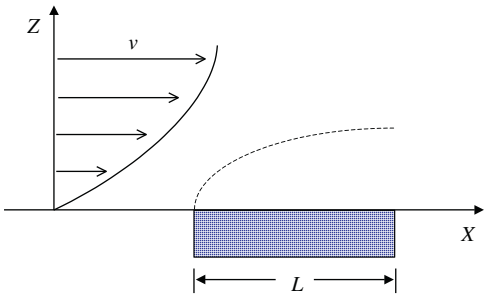


FIGURE 9.9 Convective-diffusion model for dissolution of a planar surface with liquid flow parallel to the surface. (The dashed curve indicates the boundary layer region. X and Z are coordinate axes, L is the solid surface length, and v is the bulk fluid velocity.)

parallel to the surface contribute to the dissolution of the solid. For such a flat surface, the diffusion boundary layer is well characterized and can be used to calculate the dissolution rate. An equation derived by Levich can be applied to the dissolution of a rectangular tablet of length L (in the direction of flow) and width b[25]:

$$\frac{dQ}{dt} = 0.68D^{2/3}C_s u^{-1/6} v^{1/2} b L^{1/2} \tag{9.61}$$

where:

D is the diffusion coefficient,
C_s is the solubility,
u is the kinematic viscosity,
v is the bulk fluid velocity.

Nelson and Shah also developed a convective-diffusion model for planar surface dissolution under parallel flow.[26–29] For rectangular tablets with length L (in the direction of flow) and width b, the dissolution rate is

$$\frac{dQ}{dt} = 0.808D^{2/3}C_s \alpha^{1/3} b L^{2/3} \tag{9.62}$$

where:

α is the rate of shear in the boundary layer,
b is the tablet width (normal to flow),
L is the tablet length (parallel to flow).

If the tablet is circular with a radius of r, the equation becomes

$$\frac{dQ}{dt} = 2.157D^{2/3}C_s \alpha^{1/3} r^{5/3} \tag{9.63}$$

The results obtained by Nelson and Shah were confirmed by Neervannan et al. using numerical methods to solve the convective-diffusion equations.[30]

9.2.4 Particulate dissolution

Modeling planar surface dissolution is important for understanding dissolution mechanisms. However, dissolution of solid dosage forms usually involves solid particles. Most rapid release oral dosage forms first disintegrate into small particles, which then dissolve in order for the drug to be bioavailable. The same dissolution theories that apply to planar surface dissolution are also applicable to particle dissolution. However, the dissolution of solid particles is more complicated because of surface area changes during dissolution. Also, particles have various shapes and size distributions, which make modeling their dissolution process even more complicated.

9.2.4.1 *Diffusion layer–based dissolution models*

Three diffusion-controlled models for monosized spherical particle dissolution under sink conditions are given by

$$W^{1/3} = W_0^{1/3} - K_{1/3}t \quad K_{1/3} = \left(\frac{4\pi\rho N}{3}\right)^{1/3} \frac{DC_s}{\rho h} \tag{9.64}$$

$$W^{1/2} = W_0^{1/2} - K_{1/2}t \quad K_{1/2} = \left(\frac{3\pi\rho N}{2}\right)^{1/2} \frac{DC_s}{k'\rho} \tag{9.65}$$

$$W^{2/3} = W_0^{2/3} - K_{2/3}t \quad K_{2/3} = \left(\frac{4\pi\rho N}{3}\right)^{2/3} \frac{2DC_s}{\rho} \tag{9.66}$$

where:

W is the particle weight at time t,
W_0 is the initial particle weight,
$K_{1/3}$, $K_{1/2}$, and $K_{2/3}$ are composite rate constants,
ρ is the density of the particle,
N is the number of particles,
D is the diffusion coefficient,
C_s is the solubility,
h is the diffusion layer thickness,
k' is a constant.

Eq. (9.64), reported by Hixson and Crowell, is known as the *cube-root law*.[16,31] It was derived originally based on the assumption that dissolution rate is proportional to the surface area of spherical particles. It can also be derived using a simple diffusion layer model. Eq. (9.65) is a semiempirical expression reported by Niebergall et al. and has a square-root dependency on weight.[32] Eq. (9.66) has a two-thirds-root dependency on weight and was derived by Higuchi and Hiestand for the dissolution of very small particles.[17,33]

Each of these three particle dissolution models gives a satisfactory fit to certain experimental dissolution data.[31–33] However, the choice of the model to fit experimental data is somewhat arbitrary. Although they appear different in form, the three equations are often difficult to distinguish when applied to experimental data. Pedersen and Brown evaluated these three models experimentally by studying the dissolution of tolbutamide powder.[34] By taking into account the particle size distribution in their computer model, they found that the cube-root model worked best with the smallest mean squared deviation for the fitted curve. However, the fitting results of the cube-root and the square-root models were almost equally good.

By accounting for the curvature of the concentration gradient in the diffusion layer around a spherical particle, Wang and Flanagan derived a general solution for diffusion-controlled particle dissolution (Eqs. (9.67) and (9.68)).[18] Their analysis shows that the concentration gradient inside a diffusion layer is not linear

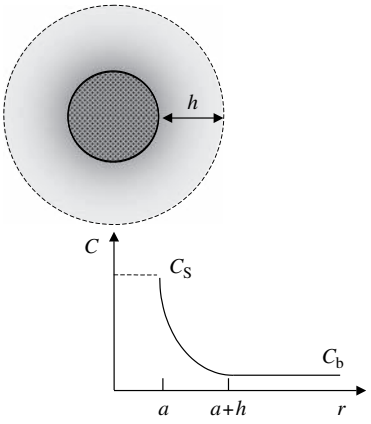


FIGURE 9.10 Pseudo-steady state concentration gradient around a spherical particle.

(Fig. 9.10). Mathematical analysis shows that the three classical particle dissolution rate expressions are approximate solutions to the general solution. The cube-root law is most appropriate when the particle size is much larger than the diffusion layer thickness, the two-thirds-root expression applies when the particle size is much smaller than the diffusion layer thickness. The square-root expression is intermediate between these two models:

$$\frac{DC_s}{\rho h}t = a_0 - a - h\ln\frac{h+a_0}{h+a} \tag{9.67}$$

$$\frac{DC_s}{\rho h}t = \left(\frac{3w_0}{4\pi\rho}\right)^{\frac{1}{3}} - \left(\frac{3w}{4\pi\rho}\right)^{\frac{1}{3}} h\cdot\ln\frac{h+\left(\frac{3w_0}{4\pi\rho}\right)^{\frac{1}{3}}}{h+\left(\frac{3w}{4\pi\rho}\right)^{\frac{1}{3}}} \tag{9.68}$$

In deriving the general solution for diffusion-controlled particle dissolution, pseudo-steady state is an important assumption. For example, dissolution from a planar surface under sink conditions reaches steady state when the dissolution rate becomes constant; hence, the concentration gradient in the diffusion layer becomes time-invariant. In any short time interval, the number of molecules moving into the diffusion layer from one side equals the number of molecules moving out of the diffusion layer from the other side.

For particle dissolution, a similar state can often be reached in which the number of molecules moving into the diffusion layer from one side approximately equals the number of molecules moving out of the diffusion layer from the other side. This assumption requires that within the time for an average molecule to cross the diffusion layer, the change of particle surface area is negligible.

An interesting result derived from the general particle dissolution model is given by Eq. 9.69, which demonstrates that the dissolution rate per unit surface area (surface-specific dissolution rates) depends upon particle size, with smaller particles having higher surface-specific dissolution rates. Bisrat and Nystrom[35] and Anderberg et al.[36,37] reported a dependence of surface-specific dissolution rates upon particle size (with radius of a). Their results showed the same trend as predicted here. The surface-specific dissolution rate increased (after correcting for solubility dependence on particle size) with decreasing particle size. This increase was especially pronounced for particle sizes below about 5 μm:

$$J = DC_s\left(\frac{1}{h}+\frac{1}{a}\right) \tag{9.69}$$

9.2.4.2 *Convective-diffusion-based particulate dissolution model*

The convective diffusion model has the advantage of being more hydrodynamically realistic. Applying convective diffusion to spherical particle dissolution has the potential of being able to predict dissolution rate more accurately.

Levich derived an equation for convective diffusion from a sphere. Assuming that liquid flow around a particle is in creeping motion while the boundary layer is small compared to particle size, Levich derived an expression for mass transport to a free-falling sphere based on convective diffusion.[21] This expression is applicable when the Reynolds number is much less than 1 while the Peclet number is much larger than 1. When this expression is applied to spherical particle dissolution, it can be written as

$$\frac{dQ}{dt} = 7.98\cdot C_s D^{2/3} v^{1/3} a^{4/3} \tag{9.70}$$

where:

Q is the amount dissolved,
t is time,
C_S is solubility,
D is the diffusion coefficient,
v is the bulk solvent velocity,
a is the particle radius.

The dissolution rate per unit area (J) is obtained by normalizing the total dissolution rate (Eq. 9.70) by particle surface area ($4\pi a^2$), which leads to

$$J = 0.635\cdot C_s D^{2/3} v^{1/3} a^{-2/3} \tag{9.71}$$

The significance of Eq. 9.71 is that the dissolution rate per unit area depends on surface curvature. For the same total surface area, smaller particles have higher dissolution rates. Although the expressions are different, this conclusion qualitatively agrees

well with the previous general diffusion layer model for spherical particle dissolution (Eq. 9.69). Applying this equation for diffusional flux under such flow conditions leads to the following equation for particle weight undissolved (w) with time:

$$w^{5/9} = w_0^{5/9} - 2.35 \cdot C_s D^{2/3} v^{1/3} \rho^{-4/9} t \tag{9.72}$$

For dissolution of N monosized spherical particles, the equation becomes

$$w^{5/9} = w_0^{5/9} - K_{5/9} t \quad K_{5/9} = 2.35 \cdot C_s D^{2/3} v^{1/3} \rho^{-4/9} N^{5/9} \tag{9.73}$$

For dissolution of a polydisperse particle system, the overall dissolution profile can be obtained by summing the contributions from each particle:

$$W_t = \sum_{i=1}^{N} W_{ti} \tag{9.74}$$

where:

W_t is the total particle weight remaining at time t,
w_{ti} is the weight remaining of particle i at time t,
N is the number of particles.

Since all the parameters in Eq. 9.73 can be determined independently, it is possible to calculate particle dissolution profiles a priori without resorting to parameters that must be calculated from the dissolution data (ie, diffusion layer thickness).

The derivation of Eq. 9.73 based on Levich's expression is a novel approach to modeling particle dissolution. Independent of the chapter authors, Abdou[37] and Shah et al.[38] presented equations similar to Eq. 9.73. Their equations were slightly different from Eq. 9.73, possibly due to editorial errors.

9.2.4.3 Dissolution under nonsink conditions

For solid dissolution, a changing bulk solution concentration complicates the modeling of this process. The assumption of sink conditions makes the mathematical derivation easier. The particle dissolution models summarized in this chapter are such examples. However, many dissolution processes take place under nonsink conditions, particularly for sparingly soluble compounds.

If the change of solvent volume (V) during dissolution is minimal, the concentration in the bulk solution can be expressed as

$$C_b = \frac{Q}{V} \tag{9.75}$$

where:

Q is the amount dissolved,
V is the solvent volume.

For spherical particle dissolution, concentration in the bulk solution is expressed as

$$C_b = \frac{\frac{4}{3}\pi\rho(a_0^3 - a^3)}{V} \tag{9.76}$$

By introducing Eqs. 9.75 and 9.76 to the abovementioned particle dissolution models, the particle dissolution models discussed here can be extended to nonsink conditions. Hixson and Crowell actually derived equations for particle dissolution under nonsink conditions. They also discussed a special situation in which the initial total particle weight is equal to the saturation capacity of the solvent. However, their equations were presented in a way that was inconvenient to apply.[16] These equations were later derived and discussed by Short et al.[39] and by Pothisiri and Carstensen.[40] Patel and Carstensen found that these nonsink condition models were valid up to 80–90% dissolved for the dissolution of p-hydroxybenzoic acid and sodium chloride.[41]

9.2.4.4 Effects of particle shape

All three classical particle dissolution models discussed in this chapter were developed for spherical particles. Obviously, real pharmaceutical powders usually have irregular shapes. It is desirable to know how much deviation irregular particle shapes can contribute to the dissolution profile. Pedersen and Brown investigated this problem theoretically.[42] Assuming exact isotropic dissolution, they derived the equations governing the dissolution profile for 10 common crystal shapes and found that the dissolution of each crystal shape can usually be approximated by a hypothetical spherical particle dissolution. The deviation became large after a large portion of the particle had dissolved.

Dali and Carstensen studied the effect of particle shape on dissolution behavior using single potassium dichromate crystals.[43] They incorporated a semiempirical shape factor to fit the dissolution profile of a crystal to that of a spherical particle of equivalent volume. Assuming that deviations from the cube-root law were completely due to the shape factor, they found that the shape factor is not constant; rather, it changes as a particle dissolves.

9.2.4.5 Polydispersity effects

Classical particle dissolution models, like the cube-root law, are intended only for single- or monodisperse-particle dissolution. However, pharmaceutical powders are composed of particles of various sizes. Sieve cuts from powders have often been used in dissolution experiments. Based on theoretical calculations, Brooke pointed out that, when treated as monosized powders, the error can be as much as 3% for the narrowest sieve cuts depending on particle size distribution across the cut. For wider sieve cuts, errors of 6–10% are possible.[44] Therefore, a theoretical treatment for polydispersity

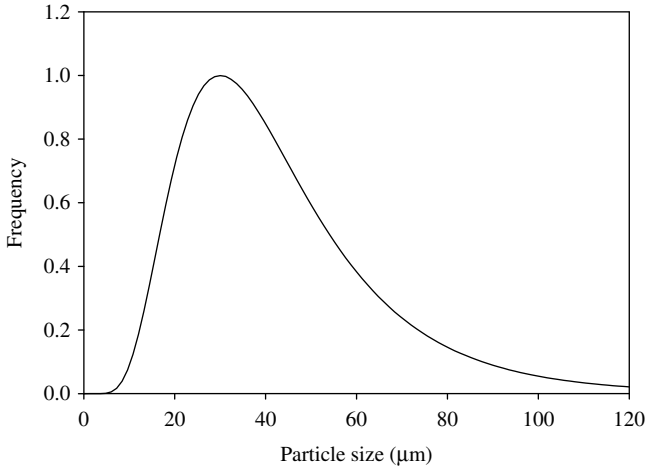


FIGURE 9.11 Typical log-normal particle size distribution for pharmaceutical powders.

effects is necessary even for fairly narrowly distributed powders. Most pharmaceutical powders are produced by milling and, as a result, tend to give skewed particle size distributions, as shown in Fig. 9.11. Such a distribution can be converted to a Gaussian distribution by plotting the x-axis on a logarithmic scale. Such a distribution is log-normal. Most theoretical investigations for polydispersed particle dissolution assume the particle size distribution to be log-normal.

Higuchi et al. studied the dissolution of a polydispersed methylprednisolone powder by approximating a log-normal distribution with a simpler function.[33] By convoluting the approximate particle size distribution function with the two-thirds-root particle dissolution model, the total particle dissolution profile was calculated. Reasonably good agreement between experimental results and theory was obtained in the study of methylprednisolone powder dissolution.

Brooke derived an exact equation for the dissolution of a log-normal distributed powder based on the cube-root single-particle dissolution model.[45] Although the calculations are tedious, they can be performed with a calculator and a normal distribution table. Brooke's expression is

$$\begin{aligned} W_\tau = {} & \mathrm{re}^{3(\mu+3\sigma^2/2)}\left(1-F\left[\frac{\ln(\tau)-(\mu+3\sigma^2)}{\sigma}\right]\right) \\ & -3\mathrm{r}\tau\mathrm{e}^{2(\mu+\sigma^2)}\left(1-F\left[\frac{\ln(\tau)-(\mu+2\sigma^2)}{\sigma}\right]\right) \\ & +3\mathrm{r}\tau^2\mathrm{e}^{(\mu+\sigma^2/2)}\left(1-F\left[\frac{\ln(\tau)-(\mu+\sigma^2)}{\sigma}\right]\right) \\ & -r\tau^3\left(1-F\left[\frac{\ln(\tau)-\mu}{\sigma}\right]\right) \end{aligned} \tag{9.77}$$

where

$$F(x)=\frac{\sqrt{2}}{2\pi}\int_{-\infty}^{x} e^{-x^2/2}dx$$

and

$$\tau=\frac{2\mathrm{k}C_\mathrm{s}}{\rho t}$$

W_τ is the weight undissolved,
μ is the mean particle diameter,
σ is its standard deviation,
$r = \pi\rho N/6$,
C_S is the particle solubility,
ρ is the particle density,
k is a constant.

Brooke also pointed out that many actual powders follow truncated distributions. Therefore, it is possible that applying an untruncated distribution or ideal log-normal distribution would lead to errors. If the exact equation is used to represent a distribution that is truncated, the result is erroneously adding some larger or smaller particles that do not actually exist. However, the error due to the addition of smaller particles is minimal, while the error due to the addition of larger particles can be significant. This arises because for the same number of particles, the total weight of the larger particles is much higher than that for smaller particles. Brooke then presented an equation for the dissolution of truncated log-normal distributed powders, which is applicable before the critical time (ie, before the smallest particles disappear).[46] Pedersen and Brown and Pedersen presented complex expressions for polydispersed powder dissolution under sink or nonsink conditions based on the three classical single-particle dissolution models.[47–49]

With faster computational tools, it is less important to have a single equation convolving a size distribution function with a particular basic particle dissolution model to predict a dissolution profile. So long as the size distribution is available, the total dissolution profile can be calculated by treating each narrow size fraction separately and adding the contributions of each fraction together. The additional advantage for this approach is that particle size does not need to follow any specific distribution. Several powder dissolution studies have been conducted using this method, and satisfactory results were obtained.[50–52]

In all of the methods dealing with polydispersed particles discussed in this chapter, it is assumed that each particle dissolves independently. This may not be true under certain experimental conditions. A commonly encountered problem in particle dissolution is aggregation. In such cases, the dissolution of one particle is affected by the existence of other particles. The

abovementioned multiparticle dissolution expressions do not apply to such situations.

References

1. Jacobs MH. *Diffusion process*. Berlin: Springer-Verlag; 1967.
2. Crank J. *The mathematics of diffusion*. 2nd ed. Oxford: Clarendon Press; 1986.
3. Berg HC. *Random walks in biology*. Princeton University Press; 1983.
4. Cussler EL. *Diffusion: mass transfer in Fluid Systems*. 2nd ed. Cambridge: Cambridge University Press; 1997.
5. Carslaw HS, Jaeger JC. *Conduction of heat in solids*. 2nd ed. Oxford: Oxford University Press; 1959.
6. Flynn GL, Yalkowsky SH, Roseman TJ. Mass transport phenomena and models: theoretical concepts. *J Pharm Sci* 1974;**63**:479–510.
7. Qiu Y, Chen Y, Zhang GZ. *Developing solid oral dosage forms*. New York: Academic Press; 2009.
8. Cooper ER. Effect of adsorption on membrane diffusion. *J Colloid Interface Sci* 1974;**48**:516–17.
9. Chen, Y. (1994). Ph.D. Thesis. University of Iowa.
10. Hunke WA, Matheson LE. Mass transport properties of co(polyether)polyurethane membranes II: permeability and sorption characteristics. *J Pharm Sci* 1981;**70**:1313–18.
11. Higuchi T. Rate of release of medicaments from ointment bases containing drugs in suspension. *J Pharm Sci* 1961;**50**:874–5.
12. Higuchi T. Mechanism of sustained-action medication. Theoretical analysis of rate of release of solid drugs dispersed in solid matrices. *J Pharm Sci* 1963;**52**:1145–9.
13. Langer R, et al. Zero-order controlled-release polymer matrices for micro- and macromolecules. *J Pharm Sci* 1983;**72**:17–22.
14. Noyes AA, Whitney WR. The rate of solution of solid substances in their own solutions. *J Am Chem Soc* 1897;**19**:930.
15. Nernst W. Theorie der reaktionsgeschwindigkeit in heterogenen systemen. *Z Physik Chem* 1904;**47**:52.
16. Hixson AW, Crowell JH. Dependence of reaction velocity upon surface and agitation I: Theoretical consideration. *Ind Eng Chem Res* 1931;**23**:923.
17. Higuchi WI, Hiestand EN. Dissolution rates of finely divided drug powders. I. Effect of a distribution of particle sizes in a diffusion-controlled process. *J Pharm Sci* 1963;**52**:67–71.
18. Wang J, Flanagan DR. General solution for diffusion-controlled dissolution of spherical particles I: Theory. *J Pharm Sci* 1999;**88**:731.
19. Bruner E. Reaktionsgeschwindigkeit in heterogenen systemen. *Z Physik Chem* 1904;**47**:56–102.
20. Levich VG. *Physicochemical hydrodynamics*. 2nd ed. Englewood Cliffs, NJ: Prentice-Hall; 1962. p. 57–60.
21. Levich VG. *Physicochemical hydrodynamics*. 2nd ed. Englewood Cliffs: Prentice-Hall; 1962. p. 61–72.
22. Grant DJW, Brittain HG. Solubility of pharmaceutical solids. In: Brittain HG, editor. *Physical characterization of pharmaceutical solids*. New York, NY: Marcel Dekker, Inc; 1955. p. 359.
23. McNamara DP, Amidon GL. Dissolution of acidic and basic compounds from the rotating disk: influence of convective diffusion and reaction. *J Pharm Sci* 1986;**75**:858–68.
24. McNamara DP, Amidon GL. Reaction plane approach for estimating the effects of buffers on the dissolution rate of acidic drugs. *J Pharm Sci* 1988;**77**:511–17.
25. Levich VG. *Physicochemical hydrodynamics*. 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, Inc; 1962. p. 87–91.
26. Nelson KG, Shah AC. Convective diffusion model for a transport-controlled dissolution rate process. *J Pharm Sci* 1975;**64**:610–14.
27. Shah AC, Nelson KG. Evaluation of a convective diffusion drug dissolution rate model. *J Pharm Sci* 1975;**64**:1518–20.
28. Nelson KG, Shah AC. Mass transport in dissolution kinetics. I: Convective diffusion to assess the role of fluid viscosity under forced flow conditions. *J Pharm Sci* 1987;**76**:799–802.
29. Shah AC, Nelson KG. Mass transport in dissolution kinetics. II: convective diffusion to assess role of viscosity under conditions of gravitational flow. *J Pharm Sci* 1987;**76**:910–13.
30. Neervannan S, et al. A numerical convective-diffusion model for dissolution of neutral compounds under laminar flow conditions. *Int J Pharm* 1993;**96**:167–74.
31. Hixson AW, Crowell JH. Dependence of reaction velocity upon surface and agitation II: Experimental procedure in study of surface. *Ind Eng Chem* 1931;**23**:1002–9.
32. Niebergall PJ, Milosovich G, Goyan JE. Dissolution rate studies. II. Dissolution of particles under conditions of rapid agitation. *J Pharm Sci* 1963;**52**:236–41.
33. Higuchi WI, Rowe EL, Hiestand EN. Dissolution rates of finely divided drug powders. II. Micronized methyl-prednisolone. *J Pharm Sci* 1963;**52**:162–4.
34. Pedersen PV, Brown KF. Experimental evaluation of three single-particle dissolution models. *J Pharm Sci* 1976;**65**:1442–7.
35. Bisrat M, Nystrom C. Physicochemical aspects of drug release VIII: The relation between particle size and surface specific dissolution rate in agitated suspensions. *Int J Pharm* 1988;**47**:223–31.
36. Anderberg EK, Bisrat M, Nystrom C. Physicochemical aspects of drug release VII: The effect of surfactant concentration and drug particle size on solubility and dissolution rate of felodipine, a sparingly soluble drug. *Int J Pharm* 1988;**47**:67–77.
37. Abdou HM. *Theory of dissolution. Dissolution, Bioavailability and Bioequivalence*. Easton, PA: Mack Publishing Company; 1989. p. 11–36.
38. Shah AC, Nelson KG, Britten NJ. *Pharm Res* 1991;**8**:S-182.
39. Short MP, Sharkey P, Rhodes CT. Dissolution of hydrocortisone. *J Pharm Sci* 1972;**61**:1732–5.
40. Pothisiri P, Carstensen JT. Nonsink dissolution rate equation. *J Pharm Sci* 1973;**62**:1468–70.
41. Patel M, Carstensen JT. Nonsink dissolution rate equations. *J Pharm Sci* 1975;**64**:1651–6.
42. Pedersen PV, Brown KF. Theoretical isotropic dissolution of nonspherical particles. *J Pharm Sci* 1976;**65**:1437–42.
43. Dali MV, Carstensen JT. Effect of change in shape factor of a single crystal on its dissolution behavior. *Pharm Res* 1996;**13**:155–62.
44. Brooke D. Sieve cuts as monodisperse powders in dissolution studies. *J Pharm Sci* 1975;**64**:1409–12.
45. Brooke D. Dissolution profile of log-normal powders: exact expression. *J Pharm Sci* 1973;**62**:795–8.
46. Brooke D. Dissolution profile of log-normal powders. II. Dissolution before critical time. *J Pharm Sci* 1974;**63**:344–6.
47. Pedersen PV, Brown KF. Dissolution profile in relation to initial particle distribution. *J Pharm Sci* 1975;**64**:1192–5.
48. Pedersen PV. New method for characterizing dissolution properties of drug powders. *J Pharm Sci* 1977;**66**:761–6.
49. Pedersen PV, Brown KF. General class of multiparticulate dissolution models. *J Pharm Sci* 1977;**66**:1435–8.
50. Mauger JW, Howard SA, Amin K. Dissolution profiles for finely divided drug suspensions. *J Pharm Sci* 1983;**72**:190–3.
51. Lu ATK, Frisella ME, Johnson KV. Dissolution modeling: factors affecting the dissolution rates of polydisperse powders. *Pharm Res* 1993;**10**:1308–14.
52. Almeida LPD, et al. Modeling dissolution of sparingly soluble multisized powders. *J Pharm Sci* 1997;**86**:726–32.

CHAPTER

10

Particle, Powder, and Compact Characterization

G.E. Amidon[1], *P.J. Meyer*[2] *and D.M. Mudie*[3]

[1]University of Michigan, Ann Arbor, MI, United States [2]Pfizer, Inc. Portage, MI, United States
[3]Bend Research, Bend, OR, United States

10.1 INTRODUCTION

As timelines become tighter, internal and external pressures abound for the pharmaceutical industry. The cost of drug development continues to rise at an alarming rate (with an average cost to market of $1–2 billion per drug), making it imperative to develop drugs in more effective ways to reduce cost and time. Historically, early drug product development has focused on using the design of experiments to manufacture numerous "kilogram size batches" to optimize the formulation, and gain a complete understanding of the manufacturing process. This method may produce a high-quality product, but requires the use of large quantities of expensive active pharmaceutical ingredient (API), and results in lengthy timelines.

Since most compounds do not make it through the development stages, a different approach is needed for product development. Hence, efforts have focused on using material-sparing strategies to develop early solid dosage forms. These materials—sparing formulations are developed by characterizing particle, powder, and compacts of API, excipients, and formulations at small-scale (50–200 g). Such characterization, coupled with predictive tools, allows scientists to understand the important physical, chemical, and mechanical properties of materials (eg, critical material attributes) to design robust formulations quickly at a relatively low cost and achieve acceptable exposure in clinical studies. Some of the characterization methods discussed in this chapter are summarized in Table 10.1.

There are advantages and disadvantages to the material-sparing paradigm. Using smaller amounts of material results in a lower cost, but may increase the risk of working with a nonrepresentative sample, which may result in errors. To reduce resource requirements, we have proposed strategies to decrease method development time and introduced some general guidelines to maximize the chances of success.

10.2 PARTICLE SIZE CHARACTERIZATION

Particle properties such as size, size distribution, shape, and texture are important considerations in the development of formulations using the material-sparing approach. This is true for API, excipients, and formulations (blends and granulations). When characterizing API, excipients, or formulations, a representative sample must be obtained. The sample needs to be correctly and reproducibly prepared, and necessary parameters must be correctly used for the analysis.[1] Each particle sizing method, replete with its assumptions, results in a unique measure of particle size and distribution since a sample exhibits a range of shapes and sizes and, of course, the complexity is compounded when analyzing multicomponent systems such as formulations. Since no two methods of particle sizing will result in the same "numbers," it becomes important to distinguish between different methods and to use the same type of analysis when comparing lots of API, excipients or formulations, examining differences in performance that process changes make, among other things. While many types of analyses for particle characterization are available and possible, this discussion focuses on those methods used in the material-sparing paradigm. Table 10.2 outlines the various techniques used for particle characterization in a material-sparing formulation development process. For those interested in more detailed information on particles size measurements, standard textbooks are recommended.[2]

DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00010-8

TABLE 10.1 Material-Sparing Characterization Methods

Method	Measured parameters
PARTICLE CHARACTERIZATION	
Light microscopy	Size, shape, roughness, size range
Polarized light microscopy	Crystallinity
Scanning electron microscopy	Size, shape, roughness, size range
Sieving	Size, size distribution
Light diffraction particle size	Quantitative size, distribution, span
POWDER CHARACTERIZATION	
Helium pycnometry	True density
Bulk/tapped density	Bulk and tapped density, compressibility index
Shear cell	Powder flow parameters, flow function coefficient, unconfined yield strength, cohesion, effective angle of internal friction
COMPACT CHARACTERIZATION	
Tablet compaction	Compaction pressure, solid fraction
Indentation test	Deformation pressure, elastic deformation
Tensile test	Tensile strength, compromised tensile strength
TABLET CHARACTERIZATION	
Tabletability	Tensile strength—solid fraction relationship
Compactibility	Tensile strength—compression pressure relationship
Compressibility	Solid fraction—compression pressure relationship
Manufacturability	Tablet crushing force—compression force relationship

TABLE 10.2 Summary of Material-Sparing Particle Characterization Techniques

Method	Min, micron	Max, microns	Distribution	Shape	Texture	Grams	Crystallinity
MICROSCOPY							
Light	1	1000	Yes	Yes	Yes	<0.1	No
Polarized	1	1000	Yes	Yes	Yes	<0.1	Yes
SEM	0.02	1000	Maybe	Yes	Yes	<0.1	No
Dynamic image analysis	1	10000	Yes	Yes	No	2–5	No
LIGHT SCATTERING							
Laser—wet	0.02	2000	Yes	No	No	1–2	No
Laser—dry	0.02	2000	Yes	No	No	1–2	No
OTHER							
Sieving	25–50	2000	Yes	No	No	3–5	No

Particle size measurements vary in difficulty depending on the shape of the particle. Spherical particles, for example, can be fully defined by their diameter. In reality, most particles are far from spherical; however, for certain analyses (such as laser light diffraction (LD), discussed later in this chapter) a solid particle is considered to approximate a sphere. Since this measurement is based on a hypothetical, not

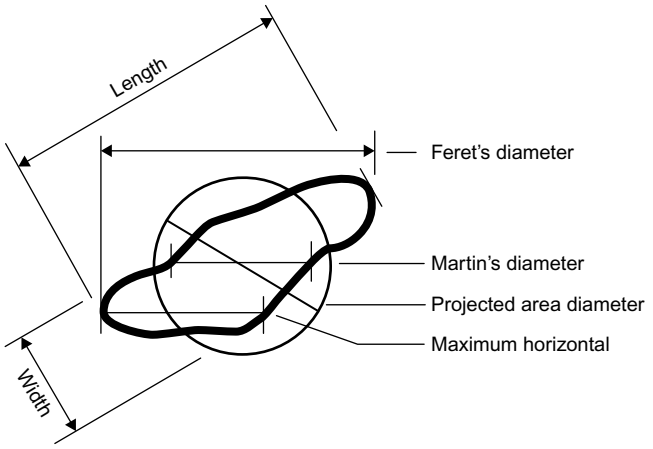


FIGURE 10.1 Particle shape parameters.

actual, sphere and is only an approximation of the true particle shape, it is referred to as the equivalent spherical diameter of the particle.[3]

For other analyses of irregularly shaped particles (such as microscopic evaluation), information about both size and shape is needed. Several different particle diameters can be used for the evaluation of particle size; the most common of these are illustrated in Fig. 10.1.[4] While length and width are diameters more often used in microscopic evaluation, the equivalent spherical diameter is widely used in laser LD analyses. Another descriptor of particle diameter used in early formulation development is the sieve diameter, which is the width of the minimum square aperture through which a particle passes.

Descriptors of particle shape include acicular, columnar, flake, plate, lath and equant, and should be included in the characterization of irregularly shaped particles. Degree of particle association (eg, aggregate, agglomerate), and surface characteristics (eg, cracked, smooth, porous), should also be noted.[4]

Since most APIs, excipients, and formulations are found in a range of sizes, size distribution of the materials is critically important and must also be determined. When a bell-shaped curve resembling a normal distribution can describe particle size distribution (PSD), it can be described by the mean particle size and standard deviation. Again, however, many pharmaceutical powders do not fall into a bell-shaped distribution. Some are skewed toward the upper end of the particle size range. These powders are often better described using a log-normal distribution that, when the data are plotted using a log scale for the size axis, resembles a normal distribution curve. When using a log-normal representation, the geometric mean standard deviation, and geometric median particle size, are used to describe the distribution. Other pharmaceutical powders cannot be characterized by a single distribution, and may have bimodal or multimodal distributions. Cumulative frequency distribution curves may also be generated by adding percent frequency values to produce a cumulative percent frequency distribution. Sieving data is often presented in this manner.

10.2.1 Light Microscopy

Light microscopy is perhaps the ultimate material-sparing characterization technique and forms the basis of the material-sparing approach as it pertains to particle characterization. With only a few milligrams of material, appearance, shape, the extent of particle association, and texture information may be quickly obtained. With a polarizing light microscope, extent of crystallinity may also be assessed. PSDs may be generated from microscopic evaluation, although it is a tedious and time-consuming process. Laser diffraction has become increasingly popular for PSD as discussed later in the chapter. More recently, dynamic image analyzers have been developed which automate the analysis of particle size and shape by capturing images with a camera, transmitting them to a computer system, and then analyzing them electronically. Average particle size, size distribution, and shape information may be generated using this technique. Image analysis may be used for particles in the size range of 1 μm−10 mm, but requires up to several grams of material depending on the difficulty of method development, and thus far has not been routinely used during the material-sparing formulation process.

For a compound light microscope, the type of microscope used, the practical size range is 1 μm−1 mm. This type of microscope is used when fine detail must be resolved. The clarity or resolving power of the image depends on the sharpness of the image produced by the objective lens. Resolving power is defined as the smallest distance between two points such that the two points can be distinguished.[5] The sample is usually mounted in a medium that has sufficient contrast to the sample, and in which the sample is insoluble. The particles should all be in the same plane, be adequately separated from each other, and should represent the size and distribution of the entire sample population (Retrieved from http://www.authorstream.com/Presentation/naveen.gokanapudi-1764246-preformulatio).

10.2.2 Scanning Electron Microscopy

Scanning electron microscopy (SEM) is another technique where only milligram quantities of material may be used to determine particle size, shape, and texture.

In SEM, a fine beam of electrons scan across the prepared sample in a series of parallel tracks. The electrons interact with the sample, and produce different signals which can be detected and displayed on the screen of a cathode ray tube (Retrieved from http://www.authorstream.com/Presentation/naveen.gokanapudi-1764246-preformulatio).[2] Particles less than 1 nm can be viewed and, since the depth of focus is so much greater than that of the light microscope, information on surface texture can be generated. SEM requires more time-consuming sample preparation than optical microscopy and cannot distinguish between crystalline and noncrystalline materials. It is also more difficult to generate a PSD using SEM since, while the information obtained is visual and descriptive, it is usually not quantitative since only a few particles are seen in the viewing field at one time. However, when SEM is used with other techniques (Retrieved from http://www.authorstream.com/Presentation/naveen.gokanapudi-1764246-preformulatio), it can provide additional valuable information that may help to explain powder properties such as agglomeration or flow problems.

10.2.3 Sieving

Sieving is a simple method that is used for determining the PSD of a powder. It is often the preferred method of choice for formulators since it is a straightforward analysis that can be done during the formulation development process (eg, after mixing or granulation of a formulation). This technique is most often used for intermediate processed materials such as granulations, formulations or excipients since larger quantities are needed (3–5 g are needed to perform an analysis using small-scale sonic sifting). API is not always evaluated by sieving due to the particle size limitations, as well as their more irregular particle shapes. Sieving is most suitable for powders whose average particle size is greater than 25–50 μm.[1]

For material-sparing sieving, a sonic sifter using 3-inch sieves is employed. Sonic sieving combines two motions to separate the particles: a vertically oscillating air column to lift particles and carry them back against mesh openings at several thousand pulses per minute, and a repetitive tapping pulse to help reduce sieve blinding (Retrieved from http://www.authorstream.com/Presentation/naveen.gokanapudi-1764246-preformulatio). Tapping also helps to deagglomerate samples with electrostatic, hygroscopic, or other adhesion problems.[6] The raw data can be converted into a cumulative weight distribution, and the data presented as a percentage of the sample retained on each sieve (histogram) or a cumulative distribution as a function of sieve size. When data is plotted on a log–log plot (sieve size vs cumulative percent), unimodal versus polymodal distributions can also be determined. A detailed procedure for analytical sieving may be found in the United States Pharmacopeia (USP) General Chapter <786>, Method I (dry sieving).[7]

10.2.4 Light diffraction

Laser LD is a preferred method for the determination of PSD of pharmaceutical materials. This technique can be quite material-sparing, often using 1–3 g of a material for a complete evaluation. Even less material may be used if combined with a microscopic or SEM evaluation to determine particle size, morphology, and fragility of the sample. There are several advantages to LD methods. Analysis time is short, robust methods can be developed with a minimum of material, results are reproducible, calibration is not required, and a wide range of measurement is possible. LD technology is divided into two general methods. High angle light scattering is appropriate for very small (submicron) particles and falls outside the scope of this discussion. Low angle light scattering is used for pharmaceutical materials in the micron size range or greater. Determination of particle size using low angle technology is only effective if one has a model to use to interpret it. For particles much larger than the wavelength of light any interaction with those particles causes the light to be scattered with only a small change in angle. This is known as Fraunhofer diffraction, and produces light intensity patterns that occur at regular intervals and are proportional to the particle diameter that produces this scatter. Fraunhofer diffraction-based instruments are applicable for particle size ranges of about 1–2000 μm, depending on the lens used.

Both wet and dry laser diffraction methods are possible and are used in a material-sparing process. For dry dispersion LD, initial conditions such as pressure and lens size are determined by microscopy, with a short series of trial runs to assess correct pressure for dispersing particles without fracturing them and to confirm lens size. Too high a pressure may break down some fragile particles. With small tray accessories, these dry methods can use 500 mg or less for a complete analysis. Initial conditions for a wet method rely on a small set of standardized generic methods and microscopy. For wet methods, it is important to choose a suspension medium where the sample has low solubility (hexane is often used) that will also disperse the sample adequately. Surfactants are occasionally used to facilitate dispersion and

inhibit flocculation of the sample, but it is important to prevent particle dissolution in the surfactant solution. A filtered saturated solution can be used as the dispersing medium to effectively prevent particle dissolution in the medium. Sonication may be needed to break up aggregates, but the fragility of the sample must be assessed since sonication may fracture individual particles and skew the results. Sample load is also important as a higher percentage of the sample in the dispersing solvent may cause aggregation.

Whether a wet or dry method is used during material-sparing formulation development depends on several factors, the most important of which are the properties of the material to be tested. Wet dispersion is frequently used for API characterization since it is a one-component system, easily screened for the appropriate nonsoluble dispersant. Multicomponent systems, such as mixes or granulations, are more amenable to dry dispersion techniques. There is no need to test the solubility of each component in the dispersant since the dispersant for dry techniques is air. Dry dispersion techniques tend to be used when there is plenty of material available, since developing a method is straightforward with sufficient trial runs. Other factors that determine which method is utilized include operator experience, equipment availability, and personal preference.

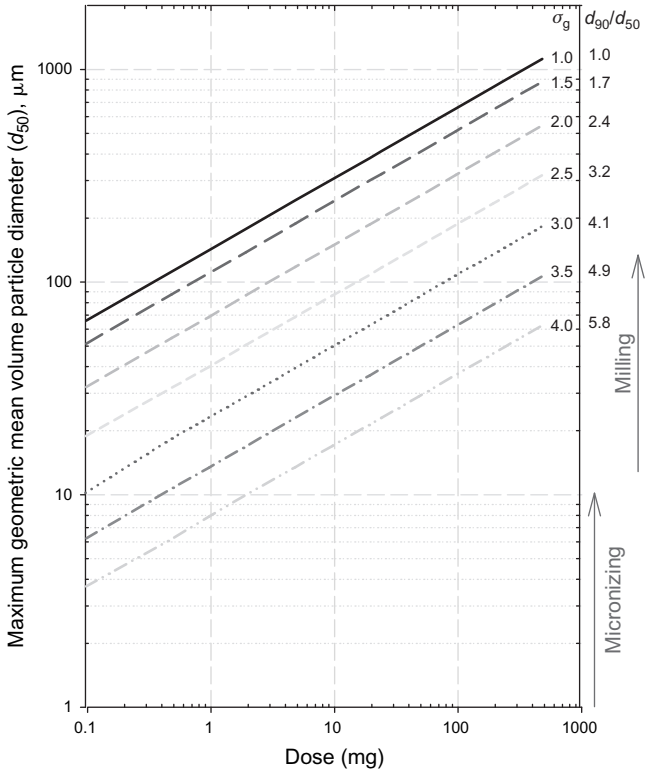


FIGURE 10.2 Content uniformity prediction as a function of particle size, distribution, and dose.

10.2.5 Importance/impact of particle size characterization

Understanding a pharmaceutical powder's particle size, shape, and distribution is an important component of material-sparing formulation development. When working with the API, a few large or small particles in a batch can alter the final tablet's content uniformity (potency, segregation), dissolution profile, and/or processing (eg, flow, compression pressure profile, and granulating properties if it is for dry granulation). Rohrs et al. have shown that with only the predicted API dose and the geometric standard deviation, an estimate of the API particle size can be obtained having a 99% probability of passing USP Stage I content uniformity criteria.[8] This is displayed graphically in Fig. 10.2. It is useful for determining approximate parameter values for particle size requirements, including whether additional API processing such as milling is necessary. API particle size and distribution data information can also help decide whether a direct compression formulation or dry granulation approach is most suitable. Examination of the API can also reveal inter- and intrabatch differences and/or trends. If for example, the PSD has changed from one batch of API to the next, this could significantly impact the processing of the final formulation.

Particle size and size distribution are also important from a dosage form performance point of view in that they are critical parameters in assuring that the desired dissolution rate is achieved for oral dosage forms. Several theoretical models for dissolution of powders have been developed.[9–11] Using a Noyes–Whitney type expression as a starting point, Higuchi and Hiestand,[10] as well as Hintz and Johnson,[11] derived expressions for dissolution of spherical particles as a function of time. While details are beyond the scope of this chapter, it is possible to predict the dissolution rate of polydispersed PSDs by summing up the predicted dissolution rate of individual size fractions of the powder. For example, using this approach, the predicted particle diameter necessary to achieve 80% dissolved in a USP dissolution apparatus in 30 minutes under sink conditions as a function of solubility and PSD is shown in Fig. 10.3.[12] With this information, selection of an appropriate particle size and distribution needed to achieve the desired dissolution rate may be estimated.

Particle characterization of the in-process or final-formulation is also critical. Flow characteristics of formulations (which will be discussed later in the chapter) are based, among other factors, on shape, size, and size distribution of particles. Large, spherical particles flow better than smaller, irregularly shaped materials, for example.

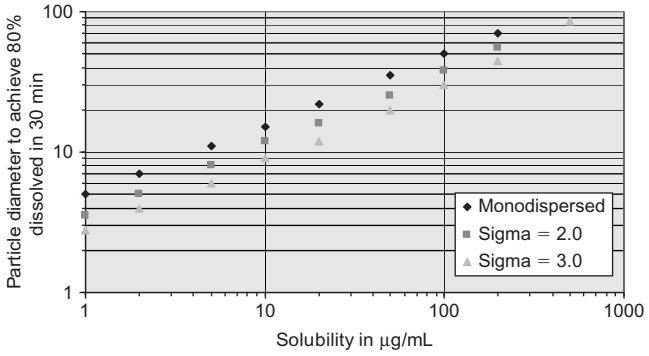


FIGURE 10.3 Predicted particle diameter necessary to achieve 80% dissolved in a USP dissolution apparatus in 30 min under sink conditions as a function of solubility and particle size distribution (sigma = geometric standard deviation of a log-normal distribution).

10.3 POWDER CHARACTERIZATION

10.3.1 Density

10.3.1.1 True density

The true density of a substance is the average mass of the particles divided by the solid volume, exclusive of all the voids that are not a fundamental part of the molecular packing arrangement; therefore, it should be independent of the method of determination.[13] There are three basic methods for determining true density:

- gas pycnometry or displacement;
- liquid displacement; and
- flotation in a liquid.

Gas pycnometry is often used in the material-sparing paradigm for determination of true density as a small amount of material is used (usually 1–8 g), the method is easy, reproducible and reliable, and the method is nondestructive: that is, the material may be reused after testing is complete.

The true density, ρ, can be calculated using Eq. (10.1).

$$\rho = \frac{w}{V_p} \tag{10.1}$$

where:

w is the weight of the sample
V_p is the powder volume.

Since the measured density is a volume-weighted average of the densities of all the individual powder particles, errors may be introduced if the gas adsorbs onto the powder or if volatile contaminants escape from the powder during the measurement. Sorption may be prevented by using an appropriate gas, usually helium or nitrogen. Volatile components may be removed during purging of the sample, and sample weight is taken before and after purging to determine if volatile contaminants were removed. True density is an essential parameter for process development and solid dosage manufacturing. As discussed in more detail later in the chapter, true density is used to calculate the solid fraction (SF) of tablets.

10.3.1.2 Bulk density

Bulk density is the mass per unit volume of a loose powder bed. The unit volume includes the spaces between the particles, and the envelope volumes of the particles themselves. The method used to fill the material into that volume can affect the degree to which the powder is compressed and can thus influence the bulk density value.[14,15] Bulk density can be calculated using Eq. (10.2), where M = mass in grams and V_o = untapped apparent volume in milliliters.

$$\text{Bulk Density } (g/\text{mL}) = \frac{M}{V_o} \tag{10.2}$$

The loose or "aerated" bulk density can be determined by allowing a defined amount of material to fill a container with a known volume under the influence of gravity.[16] The amount to which the particles collapse and fill voids between the particles depends on some powder properties, including particle shape, PSD, interparticle friction, and cohesion.

Bulk density is typically measured by gently introducing a known sample mass into a graduated cylinder, and carefully leveling the powder without compacting it. The apparent untapped volume is then read to the nearest graduated unit. As most pharmaceutical powders have densities in the range of 0.1–0.7 g/mL, a 25-mL graduated cylinder filled at least 60% full calls for a sample mass of approximately 2–11 g. (Since this test is nondestructive, the material may be reused.) USP requirements dictate a minimum graduated cylinder size of 25 mL.[16] However, if a material is in short supply, a 10-mL graduated cylinder may be used. Although wall effects could be observed, this approach provides a reasonable estimate of the bulk density.

Bulk density is an essential parameter for process development and solid dosage manufacturing. It is used in determining the amount of powder that can fit in a space such as a blender or a hopper on a tablet press or capsule filler. It is also used to determine the amount of powder that can be fitted into a capsule. Previous work has suggested that the effective bulk density of the same material will vary under different dynamics.[14,15]

10.3.1.3 Tapped density

Tapped density of a powder is the ratio of the mass of the powder to the volume occupied by the powder

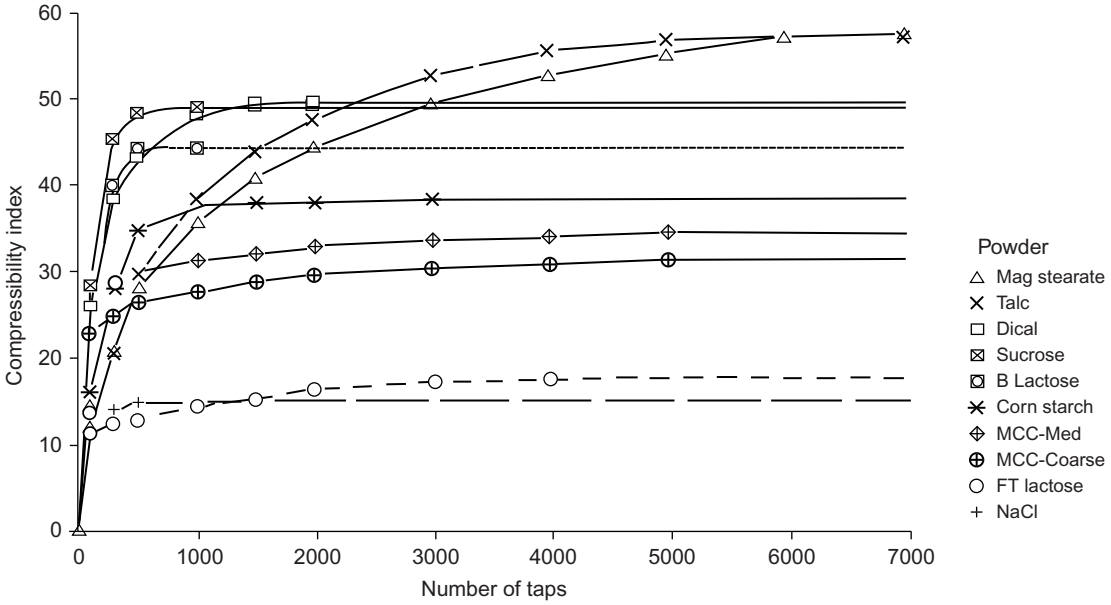


FIGURE 10.4 Influence of number of taps on the compressibility index.

after it has been tapped for a defined period of time. The tapped density of a powder represents its random dense packing. Tapped density can be calculated using Eq. (10.3), where M = mass in grams, and V_f = the tapped volume in milliliters.

$$\text{Tapped Density } (g/\text{mL}) = \frac{M}{V_f} \tag{10.3}$$

Tapped density values are higher for more regularly shaped particles (ie, spheres), as compared to irregularly shaped particles such as needles. PSD has been shown to affect the packing properties of fine powders.[17] The packing properties of a powder can affect operations critical to solid dosage manufacturing, including bulk storage, powder flow, and compaction.

Tapped density is measured by first gently introducing a known sample mass into a graduated cylinder and carefully leveling the powder without compacting it. The cylinder is then mechanically tapped by raising the cylinder and allowing it to drop under its own weight using a suitable mechanical tapped density tester that provides a suitable fixed drop distance and nominal drop rate (Retrieved from http://www.authorstream.com/Presentation/naveen.gokanapudi-1764246-preformulatio). There are two methods for measuring tapped density in the USP with different drop distance and drop rates.[16] The tap density of some pharmaceutical materials is shown in Fig. 10.4.[18] As with bulk density measurements; a material-sparing approach can be undertaken by using a 10-mL graduated cylinder (1−4 g sample requirement). This test is also nondestructive.

10.3.2 Flow

Flow assessment of APIs, excipients, and formulations are routinely completed as part of solid dosage form development. Assessments must be made to ensure powder flows adequately through processing equipment such as a roller compactor, hopper, or tablet press. Poor flowability can lead to the inability to feed powder into the dies of a rotary tablet press and can also cause tablet weight variation.

Due to the complexity of powder flow and the factors that influence it, no single measure is currently adequate for defining a flow. Unsurprisingly, many ways to measure flow currently exist, ranging from simple, qualitative methods, to more quantitative methods utilizing specialized technology. Factors such as the relative humidity of the environment, previous storage conditions, and degree of consolidation have a large impact on flowability, any of which can alter the test results.

The influence of physical and mechanical properties on powder flow is a subject of interest to formulation scientists. Factors such as PSD and particle shape have been shown to influence flow.[19–21] Additional properties such as bulk and tapped density, bonding index (BI), and internal friction coefficient are also thought to be contributors. An understanding of the effects of physical and mechanical properties on powder flowability can decrease the need to perform powder flowability analysis on some materials, resulting in significant time and resource savings.

An attempt has been made to model flowability based on physical and mechanical properties using

complex methods such as artificial neural networks, discrete element method, and constitutive models.[22–24] While these models have demonstrated a correlation between certain physical properties and the results of various methods for measuring flow, more work is required.

10.3.2.1 Compressibility Index and Hausner ratio

The compressibility index (CI) is a measure of the propensity of a powder to consolidate.[25] As such, it is a measure of the relative importance of interparticulate interactions. In a free-flowing powder, such interactions are less significant, and the bulk and tapped densities becomes closer in value. For poorer flowing materials, there are frequently greater interparticle interactions; bridging between particles often results in lower bulk density and a greater difference between the bulk and tapped densities. These differences in particle interactions are reflected in the CI. A general scale of powder flow using the CI is given in Table 10.3.[25] CI can be calculated as shown in Eq. (10.4), where V_o = untapped apparent volume, V_f = tapped apparent volume.

$$\mathrm{CI}(\%) = 100 \cdot \frac{(V_o - V_f)}{V_o} \tag{10.4}$$

Although this method cannot be used as a sole measure of powder flowability, it has the advantage of being simple to perform and it provides a quick comparison between API, excipients, and formulations. If bulk and tapped density measurements have already been performed, no additional material or experimentation is required to calculate the CI. CI has been correlated to manufacturing performance on machines such as capsule fillers. Podczeck et al. for example demonstrated a correlation between the minimum coefficient of fill weight variation and CI.[26]

The Hausner ratio (HR) is closely related to CI.[27] It can be calculated using Eq. (10.5), where V_o = untapped apparent volume and V_f = tapped apparent volume.

$$\text{Hausner Ratio} = \frac{V_o}{V_f} \tag{10.5}$$

Scales of flowability for CI and HR are included in Table 10.3.

TABLE 10.3 Scale of Flowability for Compressibility Index and Hausner Ratio[31]

Flow character	Compressibility index	Hausner ratio
Excellent	≤10	1.00–1.11
Good	11–15	1.12–1.18
Fair	16–20	1.19–1.25
Passable	21–25	1.26–1.34
Poor	26–31	1.35–1.45
Very poor	32–37	1.46–1.59
Very, very poor	>38	>1.60

10.3.2.2 Angle of repose and flow through an orifice

The angle of repose has long been used to characterize bulk solids.[28–30] Angle of repose is a characteristic related to interparticulate friction or resistance to movement between particles. According to the USP, it is the constant, three-dimensional angle (relative to the horizontal base) assumed by a cone-like pile of material formed by any of several different methods. Due to the high dependence of angle of repose measurements on testing conditions, angle of repose is not a very robust means of quantifying powder flow.[31]

Flow rate through an orifice is generally measured as the mass of material per unit time flowing from any of a number of types of containers (cylinders, funnels, hoppers). It is thought to be a more direct measure of flow than measurements such as angle of repose or HR, because it more closely simulates flow of material from processing equipment such as from a tablet press hopper into a die. Measurement of the flow rate is heavily dependent on test set-up, such as orifice diameter.[32] Both angle of repose and flow rate through an orifice methods require 5–70 g of material, and therefore are not aligned with material-sparing strategies.

10.3.2.3 Shear cell methods

Shear cell methods measure flow on a more fundamental basis than the simple methods discussed above, providing more robust flow results. Shear cell methods allow the assessment of flow properties as a function of consolidation load and time, as well as powder–hopper material interactions.[33] Various types of shear cells exist, including rotational and translational cells.[34–36] While shear cell measurements are more time-consuming than the methods discussed above, they offer a higher degree of experimental control, leading to more reproducible results. They are used extensively in multiple industries, and significant advances in automation have occurred in the past 5 years.[28,37,38]

Detailed descriptions of shear cell methods and data analysis are available in the literature.[39] The shear cell analysis method is as follows: a bulk solid sample is carefully loaded into a container or "cell," taking care not to consolidate the powder during loading. The

powder bed is typically preconsolidated at a defined load normal to the powder bed. As the load is applied, the powder bed is sheared until a uniform state of consolidation is reached (ie, powder starts to "flow" and shear stress reaches a constant value). This value of the shear stress represents the shear strength of the powder at those conditions. Once preconsolidation is achieved, the normal load is reduced, and a normal load less than the preconsolidation load is applied. The bulk solid is sheared until the shear force goes through a maximum value. This maximum value represents the shear strength of the powder bed under those conditions. This process of preconsolidation, and subsequent consolidation at a reduced load is repeated multiple times at different normal loads. This process is sometimes referred to as a yield locus test. A yield locus is constructed based on the shear data; it passes through the points defined by the shear and normal stress values during preshear and shear. Fig. 10.5 shows the yield loci of three different materials, constructed using a simplified (translational) shear cell. The yield locus is characteristic of the physical and mechanical properties of the powder, as well as a number of environmental factors (a list of some of these factors is included in Table 10.4). Some of the most commonly reported shear cell flow properties reported in the literature are the angle of internal friction and the flow function.[40,41]

Because shear methods provide quantitative, reproducible results, the properties defined in Table 10.5 can be used to compare the flowability of different drug substances, excipients, and formulations. To properly compare data, the environmental factors listed in Table 10.4 should be kept constant. For example, Moreno-Atanasio et al. demonstrated an increase in the unconfined yield strength with an increase in preconsolidation stress using a uniaxial compression test. Shear cell methods are also used to characterize the performance of materials on equipment such as capsule-filling machines,[42] and tablet presses.[43] They are extensively used to design hoppers for feeding powders, and to study the effect of storage time on flowability.[44,45] Jenike developed a mathematical methodology for determining the minimum hopper angle and opening size for mass flow from conical and wedge-shaped hoppers.[46]

Several different types of shear cell testers are available. The Schulze Ring Shear Tester is an annular or rotational shear tester. The bulk solid sample is contained in an annular shear cell containing a ring at the bottom and a baffled lid. Ring shear testers are considered to be relatively easy to operate, and provide good reproducibility.[40] A single yield locus test can be performed in approximately 30 minutes (of which the operator must be present about one-third of the time).

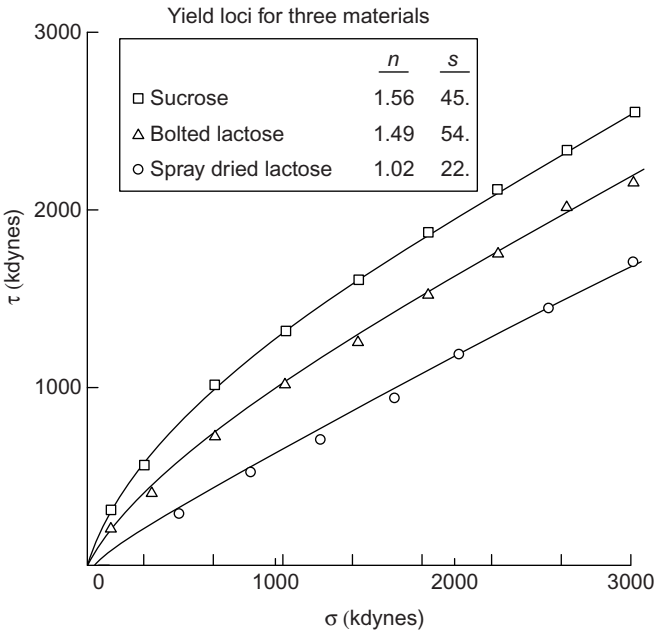


FIGURE 10.5 Yield loci for three materials.

TABLE 10.4 Physical and Mechanical Properties and Environmental Factors Influencing Flowability

Physical and mechanical properties	Environmental effects
Particle size distribution	Consolidation load
Particle shape	Consolidation time
Bulk/tapped density	Shear effects (direction & rate)
Particle density	Particle/wall interactions
Moisture content	Air permeability
Crystallinity	Compressibility (a function of load and time)
Polymorphic form	Humidity/powder moisture content
Surface area	Packed density
Electrostatic charge	
Surface energy	
Elasticity	
Plasticity (ductility)	
Viscoelasticity	
Brittleness	

Multiple shear cell sizes are available. Depending on the particle size of the powder specimen a small, 10-mL shear cell can be used, allowing for a significant reduction in sample requirements. Medium and large shear

TABLE 10.5 Shear Cell Flow Property Names and Definitions

Flow property	Definition
Major consolidation stress (σ_1)	Defined by the major principal stress of the larger Mohr stress circle
Unconfined yield strength/stress (σ_c)	Defined by the major principle stress of the smaller Mohr stress circle. It is the stress at which the sample will break (flow) after a vertical stress has been applied to a consolidated sample
Flow function coefficient (*FFC*)	The ratio of the major consolidation stress to the unconfined yield strength at a defined normal load
Flow function	A flow function curve can be constructed by plotting the FFC values obtained by performing multiple yield locus tests at different preconsolidation normal loads
Slope angle of the linearized yield locus (φ_{lin})	The angle defined by the linearized yield locus. The linearized yield locus is the line that is tangential to both Mohr stress circles
Effective angle of internal friction (φ_e)	The angle defined by the line that runs through the origin of the diagram, and is tangent to the larger Mohr stress circle
Angle of internal friction at steady state flow (φ_{sf})	The arctangent of the ratio of shear stress to normal stress of the preshear point (ie, steady state flow). It characterizes the internal friction at steady state flow in the shear plane
Cohesion (τ_c)	The shear stress at yield under zero normal stress, that is, the intersection of the yield locus with the ordinate

cells are also available, which can accommodate samples with larger particle sizes. These cell sizes require sample volumes greater than approximately 31 and 73 mL, respectively.[47] A standard test method for using the Schulze Ring Shear Tester can be found in ASTM Method D 6773.[48]

The Jenike Shear Tester is a translational tester. The shear cell is cylindrical and is split horizontally, forming a shear plane between the lower stationary base and the upper movable portion. It consists of a closed ring at the bottom, a ring of the same diameter lying above the bottom ring, and a lid.[49] An advantage of the Jenike cell over the ring shear tester is that the powder bed is shared more uniformly.[49] Disadvantages of the Jenike Shear Tester include the amount of material required and the time required for a test. Depending on the powder, and the operator's skill, 1–2 hours per yield locus (during which the operator has to be present) are required. Results generated using the Jenike Shear Tester have been shown to correspond to those generated using the ring shear tester.[39] A standard test method for using the Jenike Shear Tester can be found in the ASTM Method D 6128.[50]

10.3.2.4 Additional shear testers

Schulze ring shear and Jenike shear cell testers are used extensively due to their commercial availability, though other shear cells are available. Plate or "simplified" shear cells have also been designed, which consist of a thin sandwich of powder between a lower stationary rough surface and an rough upper surface that is movable.[18,51] Uniaxial, biaxial, and triaxial testers have also been used for flow analysis, and have been discussed in the literature. The measurement principle of the uniaxial tester is similar to that of shear cells.[35,49]

10.3.2.5 Dynamic test methods

Avalanche testers assess the flowability of powders by measuring their avalanching behavior, which is related to powder cohesivity and flowability. Unlike shear cell methodology, this type of assessment is dynamic in nature,[52] which may be more applicable to low-shear processes such as blending, in which avalanching behavior of powder promotes mixing. Avalanche testing can be carried out in different types of equipment, including rotating drums and vibratory feeders.[53,54] Avalanche testing has been shown to distinguish between freely flowing powders, blends, and granulations.[55,56]

One limitation of the avalanche tester is the qualitative nature of determining the regimes of avalanche flow in the rotating drum. Boothroyd et al. argued that the ideal flow regimes necessary for meaningful data analysis for pharmaceutical powders were the "rolling" and "cascading" regimes.[57,58] Another limitation of the system is the method development required before analysis. Rotational speed, measurement duration, and sample size must be optimized, as they have a great degree of influence over the measurement results.[53] A third limitation is the amount of material required, which limits the use of this technique in a material-sparing approach to formulation development. Hancock et al. proposed a sample size of 50 mL.[53]

Bhattachar et al. developed a vibratory feeder method for assessing avalanche behavior that requires a smaller sample size than for the Aeroflow (1.2 g). Results compared to those generated using the

Aeroflow.[59] Despite the small sample size requirements, this method is not widely used. The instrument is not commercially available and has not been extensively tested.

10.4 COMPACT (MECHANICAL PROPERTY) CHARACTERIZATION

Many investigations have demonstrated the importance and impact of the physical and chemical properties of materials on powder processing. Physical properties such as particle size and shape clearly influence powder flow for example. The previous sections of this chapter provide some recommendations for how to proceed with characterization using limited quantities of materials. However, mechanical compact properties (ie, those properties of a material under the influence of an applied stress) are also of great importance for solid dosage form development and manufacturing, particularly for tablet formulation. This section describes the importance of the mechanical properties of materials, as well as some basic principles and methodologies that can be used to investigate the influence of these properties on compaction. For this discussion, physical properties are considered to be those properties that are "perceptible especially through the senses" (ie, properties such as particle size, and shape). In contrast, mechanical properties are those properties of a material under an applied load: elasticity, plasticity, viscoelasticity, bonding, and brittleness.

Table 10.4 lists some of the physical and mechanical properties that influence powder properties and compaction. For example, surface energy and elastic deformation properties influence individual particle true areas of contact. Plastic deformation likely occurs to some extent in powder beds depending on the applied load, and almost certainly it occurs during the compaction of powders into tablets. Certainly at asperities, local regions of high pressure can lead to localized plastic yielding. Electrostatic forces can also play a role in powder flow, depending on the insulating characteristics of the material and environmental conditions. Particle size, shape, and size distribution have also been shown to influence flow and compaction. A number of environmental factors such as humidity, adsorbed impurities (air, water, among others), consolidation load and time, direction and rate of shear, and storage container properties are also important. With so many variables, it is not surprising that a wide variety of methods have been developed to characterize materials. The focus of this chapter is on those useful methods that require limited amounts of material (bulk drug or formulation) and provide the most valuable information.

What holds particles together in a tablet? A detailed discussion is beyond the scope of this chapter, and excellent references are available in the literature.[60,61] However, it is important to realize that the forces that hold particles together in a tablet or powder bed are the very same forces discussed in detail in introductory physical chemistry texts. There is nothing magical about particle–particle interactions; the forces involved are London dispersion forces, dipole interactions, surface energy considerations, and hydrogen bonding. The consolidation of powders brings particles into proximity, and these fundamental forces can begin to act effectively to produce strong particle–particle interactions (eg, bonding). Particle rearrangement, elastic and plastic deformation of material can establish large areas of true contact between particles; if the resulting particle–particle bonds are strong, a strong and intact tablet is produced.

10.4.1 Important mechanical properties

Materials used in the pharmaceutical industry can be elastic, plastic, viscoelastic, hard, tough, or brittle in the same sense that metals, plastics, or wood are. The same concepts that mechanical engineers use to explain or characterize tensile, compressive, or shear strength are relevant to pharmaceutical materials. These mechanical properties of materials can have a profound effect on solids processing.

The mechanical properties of a material play an important role in powder flow and compaction. These properties are critical properties that influence the true areas of contact between particles. Therefore, it is essential to characterize the properties. Reliable mechanical property information can be useful in helping to choose a processing method such as granulation or direct compression, selecting excipients with properties that will mask the poor properties of the drug or helping to document what went wrong, for example, when a tableting process is being scaled-up or when a new bulk drug process is being tested. Since all of these can influence the quality of the final product, it is to the formulator's advantage to understand the importance of the mechanical properties of the active and inactive ingredients and to be able to quantify the properties.

10.4.1.1 Elastic deformation

In general, during the initial stages of deformation, a material is deformed elastically. A change in shape caused by the applied stress is completely reversible, and the specimen will return to its original shape on

release of the applied stress. During elastic deformation, the stress–strain relationship for a specimen is described by Hooke's law (Eq. 10.6):

$$\sigma = E \cdot \varepsilon \tag{10.6}$$

where:

E is referred to as Young's modulus of elasticity
σ is the applied stress
ε is the strain ($\varepsilon = (l - l_o)/l_o$).

The region of elastic deformation of a specimen is shown graphically in Fig. 10.6. The reader is directed to standard texts in material science and engineering for detailed discussions of elastic deformation. As long as the elastic limit is not exceeded, only elastic deformation occurs.

The elastic properties of materials can be understood by considering the attractive and repulsive forces between atoms and molecules. Elastic strain results from a change in the intermolecular spacing and, at least for small deformations, is reversible.

10.4.1.2 Plastic deformation

Plastic deformation is the permanent change in shape of a specimen due to applied stress. The onset of plastic deformation is seen as curvature in the stress–strain curve shown in Fig. 10.7. Plastic deformation is important because it "allows" pharmaceutical excipients and drugs to establish large true areas of contact during compaction that can remain on decompression. In this way, strong tablets can be prepared.

Plastic deformation, unlike elastic deformation, is not accurately predicted from atomic or molecular properties. Rather, plastic deformation is often determined by the presence of crystal defects such as dislocations, grain boundaries, and slip planes within crystals. While it is not the purpose of this chapter to discuss this in detail, it is important to realize that dislocations and grain boundaries are influenced by factors such as the rate of crystallization, particle size, the presence of impurities, and the type of crystallization solvent used. Slip planes may exist within crystals due to molecular packing arrangements that result in weak interplanar forces. Processes that influence these (eg, crystallization rate, solvent, temperature) can be expected to influence the plastic deformation properties of materials, and hence the processing properties. The reader is directed to standard texts in material science and engineering for detailed discussions of plastic deformation.

The plastic properties of a material are often determined by an indentation test.[62] Both static and dynamic test methods are available, but all generally determine the pressure necessary to cause permanent and nonrecoverable deformation.

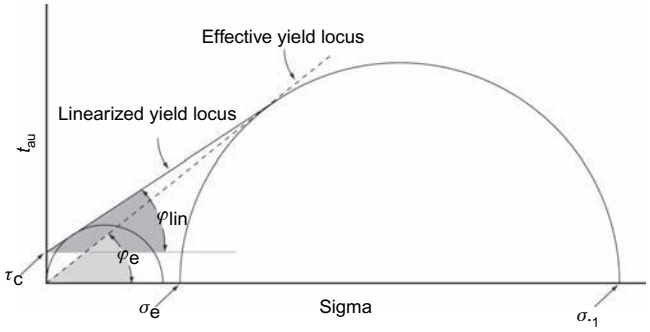


FIGURE 10.6 Yield locus and Mohr stress circle diagram.

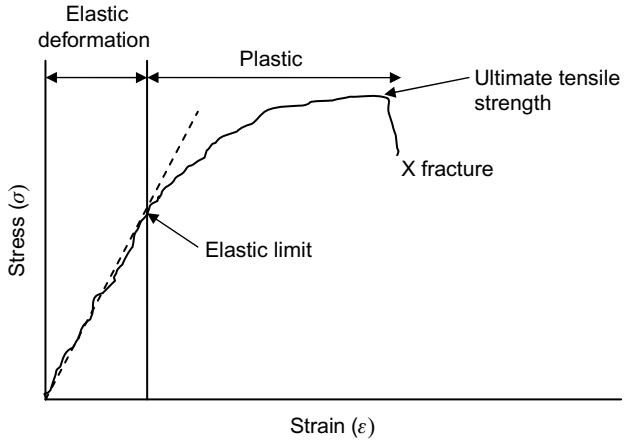


FIGURE 10.7 Stress–strain curve.

10.4.1.3 Brittle and ductile fracture

In addition to plastic deformation, materials may fail by either brittle fracture or ductile fracture; fracture being the separation of a body into two or more parts. Brittle fracture occurs by the rapid propagation of a crack throughout the specimen. Conversely, the ductile fracture is characterized by extensive plastic deformation followed by fracture. Ductile failure is not typically seen with compacts of pharmaceutical materials. The characteristic snap of a tablet during hardness testing is indicative of brittle fracture.

10.4.1.4 Viscoelastic properties

Viscoelastic properties can be important; viscoelasticity reflects the time-dependent nature of stress–strain. A basic understanding of viscoelasticity can be gained by considering processes that occur at a molecular level when a material is under stress. An applied stress, even when in the elastic region, effectively moves atoms or molecules from their energy equilibrium state. With time, the rearrangement of atoms or molecules can occur.

The stress–strain relationship can, therefore, depend on the time frame over which the test is conducted. In compacting tablets, for example, it is frequently noted that higher compaction forces are required to make a tablet with a given strength when the compaction speed is rapid. All pharmaceutical materials are viscoelastic; the degree to which their mechanical properties are influenced by rate depends on the material.

10.4.2 Overview of methods

Characterizing mechanical properties has been an active area of pharmaceutical research for decades. The application of classic "engineering" methodologies to characterize pharmaceutical materials dates to the 1950s or before. With the advent of high-speed computer control, and monitoring of processes such as tablet compaction, the era of "dynamic" characterization of pharmaceutical materials was ushered in. Sophisticated instrumentation of rotary tablet presses and, in particular, the design of tablet compaction simulators with seemingly infinite control of the compaction process, has offered scientists an unprecedented opportunity to study the mechanics of materials at speeds representative of production tablet compaction. Even today, both dynamic testing and the classic "quasi-static" engineering testing approaches offer opportunities to understand pharmaceutical materials. In this regard, dynamic and quasi-static testing are complementary tools. Both quasi-static and dynamic test methodologies will be discussed in the following sections. One key advantage of quasi-static testing is the ability to "independently" dissect out and investigate the various mechanical properties of a material. As aforementioned, pharmaceutical materials can be elastic, plastic, viscoelastic, hard, tough, or brittle. Ultimately, these individual components that cumulatively describe a pharmaceutical material determine its compaction properties in a dynamic compaction process.

The consolidation of powders into intact tablets is a process of reducing pores in a powder bed while creating interparticle bonds. During compression, materials experience complex stresses, the structure of the powder bed changes, and consolidation is brought about mainly by particle rearrangement, plastic deformation, and fragmentation.[63] The deformation of pharmaceutical materials is time dependent, and this dependency is related to the consolidation mechanism and dynamics of the consolidation process.[64–68] Under compression, for example, brittle materials are considered to consolidate predominantly by fragmentation; plastic materials deform by plastic flow. The time dependency of this process arises from stress relaxation for materials undergoing primarily plastic deformation. However, the compaction of brittle materials is often less influenced by speed because fragmentation is rapidly achieved and prolonged exposure to the force has a limited effect on tablet properties.

Several researchers have previously identified the utility of SF in describing tablet properties. Armstrong and Palfrey[69] concluded that differences in the tensile strength of tablets compressed at different speeds could be accounted for by differences in tablet porosity. Hancock and coworkers[70] found that tablet strength and disintegration time for tablets made on an eccentric press and a rotary press were comparable when considering a comparable SF. Maarschalk and coworkers[71] found that tablet tensile strength of sorbitol as a function of tablet porosity was independent of compression speed. Finally, Tye and coworkers[63] extended this work to show that tablet SF was the primary factor determining tablet strength for several pharmaceutical excipients (both brittle and ductile) over an extremely wide range of compaction speeds (dwell times from $<$10 milliseconds to 90 seconds).

The SF of a compact can be calculated based on the true density (ρ_{true}) of the material (typically determined using pycnometry), the tablet volume (ν), and the tablet weight (Wt) (Eq. 10.7):

$$SF = \frac{Wt}{\rho_{true} \cdot \nu} \tag{10.7}$$

The relationship between the SF, also referred to as relative density, and porosity (ε) is:

$$\varepsilon = 1 - SF \tag{10.8}$$

10.4.3 Quasi-static testing

Quasi-static testing typically applies variations of traditional engineering and material science testing methods to compacts (ie, test specimens) of pharmaceutical materials. There are, for example, a number of variations of indentation, tensile, flexural, compression, and brittle fracture tests in the pharmaceutical literature.[72] The quantity of material required for testing varies from 1 to 100 g. Methods for characterizing the elastic, plastic, and brittle properties of compacts of organic materials have, for example, been developed by Hiestand and coworkers.[73–77] These measures of tableting performance assess several key mechanical properties of compacted materials that have been shown to relate to tableting. Currently available methods typically require from 10 to 60 g for complete mechanical property characterization using Hiestand's methods.

10.4.3.1 Test specimen preparation

It is important to prepare test specimens of pharmaceutical materials properly so that quasi-static test results are not improperly influenced by "flaws" that may exist in the test specimen itself. Of the methods defined in the literature, the most refined method is to make square compacts using triaxial compression and decompression.[73] A split die (Fig. 10.8) is used to make compacts that are substantially free of defects that may occur if a conventional compaction process were to be used. The split die permits triaxial decompression such that the pressure applied to all three axes is essentially equal during the decompression process.[73] This is achieved by computer control of the decompression process. The stresses in the compact are more uniformly relieved in three dimensions, and this minimizes the production and propagation of flaws within the compact. Using this approach, defect free compacts of essentially any pharmaceutical material is possible, something not obtainable with uniaxial decompression.

10.4.3.2 Importance of the solid fraction

It is imperative to realize and address the fact that the mechanical properties of a compact are very much influenced by SF. A change in SF of 0.01 (ie, a change in *SF* from 0.85 to 0.86) can result in a mechanical property change of 10–20%. For this reason, it is critical to compare the properties of a material at a "reference" SF to ensure that one is "comparing apples to apples." Hiestand and coworkers[73] defined their reference SF as 0.9 (ie, porosity = 0.1% or 10%) while others have used a SF of 0.85 or even extrapolated to a SF of 1.0 (eg, zero porosity). In comparing results from the literature, it is important to keep this in mind. It is recommended that a SF in the range typical of tablet compaction be used. For compacts of organic materials, a reference SF of 0.85 is in the mid-range of those typically observed. For inorganic materials (such as dicalcium phosphate) SFs in the 0.6–0.75 range are often observed for tablets. The wide range of mechanical properties observed means no ideal value can be identified for all materials, but 0.85 is often appropriate for materials of organic origin.

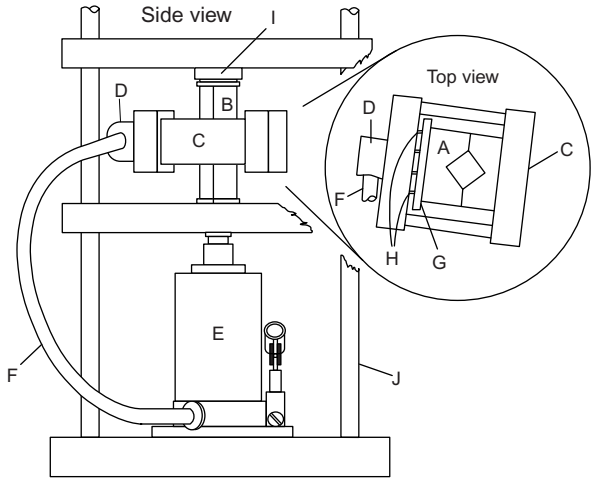


FIGURE 10.8 Schematic drawing of a simple triaxial press with a split die.

10.4.3.3 Tensile strength determination

The tensile strength, σ_T, of a square test specimen provides extremely useful information. Several methods for determining it are available and include traditional tablet hardness testing and transverse compression or square compacts. In transverse compression, specimens are compressed with platens 0.4 times the width of the compacts in the tensile testing apparatus.[73] The force necessary to cause tensile failure (tensile forces are maximum at the center of the tablet) is monitored by a load cell, and the magnitude of the force at fracture is determined. Testing of square compacts has advantages over the testing of circular compacts; however, circular compacts can be used. Conventional hardness testing of tablets can result in a measurement of tensile strength.[78] Similar results are obtained for round and square compacts when tensile failure is achieved. It is extremely important to compare measured properties such as tensile strength at the same SF. Tensile strength values more than 1 MPa (typical range 0.1–4 MPa) are typically desired for tablets.

10.4.3.4 Pendulum Impact Device

A simple schematic of a pendulum impact device (PID) is given in Fig. 10.9. This equipment permits the permanent deformation pressure of a compact of material to be determined under dynamic conditions.[73,74] Flat-faced, square tablets of the test substance are compressed at different compression pressures and then subjected to impact with a stainless steel ball in the PID. The rebound height of the ball and the chordal radius of the dent are carefully measured and used to calculate the permanent deformation pressure. In a simple sense, one is measuring the energy necessary to make the permanent deformation (the difference between the initial height of the ball and the rebound height). By measuring the volume of the dent, one can calculate the deformation pressure—the energy divided by the volume. The permanent deformation pressure is the pressure (ie, stress) necessary to cause plastic deformation. This permanent deformation pressure, *H*, has been shown to be related to the yield pressure obtained using dynamic testing methods and

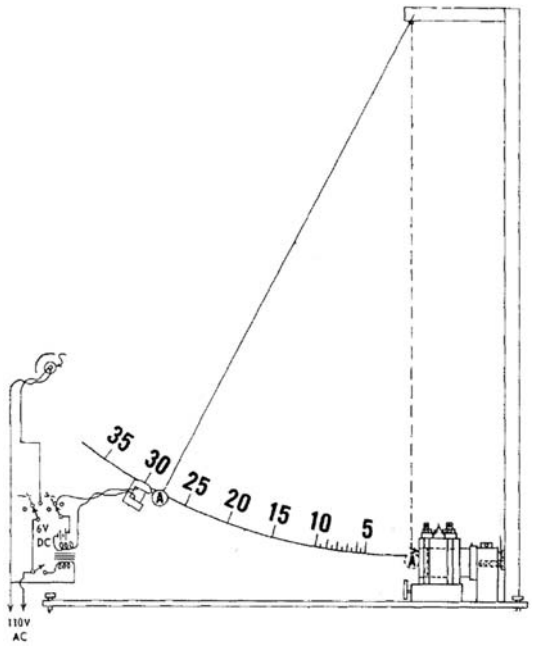


FIGURE 10.9 Pendulum impact device.

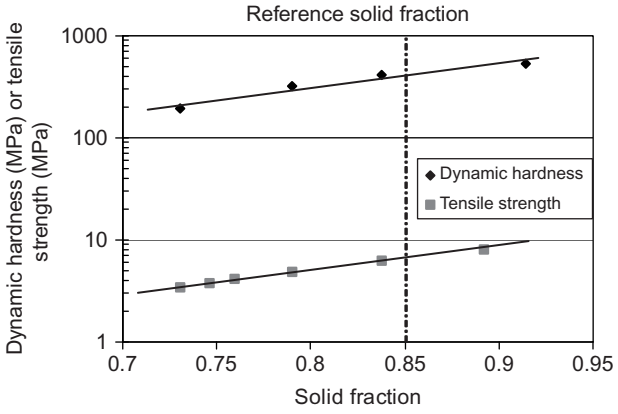


FIGURE 10.10 Dynamic hardness and tensile strength as a function of solid fraction.

Heckel analysis.[79] The dynamic hardness and tensile strength are shown in Fig. 10.10 as a function of SF for a common pharmaceutical excipient. One can clearly see the impact of SF on the measured mechanical properties.

10.4.3.5 Tableting indices

Using the methodology described above, several indices of tableting performance have been developed by Hiestand and coworkers.[60,75] These indices provide relative measures of properties (ie, dimensionless numbers) that reflect the performance of materials during processing.

10.4.3.6 Bonding Index

The purpose of the BI is to estimate the survival of strength during decompression[75]; it is defined in Eq. (10.9):

$$\mathrm{BI} = \frac{\sigma_T}{H} \tag{10.9}$$

where:

σ_T is the tensile strength of the compact at a given SF (typically 0.85 or 0.9 as defined by the user)
H is the permanent deformation pressure (ie, hardness) of a compact at the same SF.

The BI is, in essence, a measure of the ability of a material, on decompression, to maintain a high fraction of the bond that was created during compression. At maximum compression pressure, the bonded areas are at a maximum, because the right areas of contact are maximized. During decompression, some of that area and bond is "lost" due to elastic recovery. A high BI indicates that relatively speaking, a larger portion of the strength remained intact after decompression. A low BI indicates that less of the strength remains. The term BI, then, is a good description since it, in effect, characterizes the tendency of the material to remain intact after it has been compressed. Tablets made of materials with poor bonding characteristics may be quite friable. Compacts made of materials with good bonding indices may, conversely, make strong tablets. A BI in excess of 0.01 (range 0.001−0.06) is typically desired.

10.4.3.7 Brittle Fracture Index

The brittle fracture index (BFI) is a measure of the brittleness of a material. It is a measure of the ability of a compact to relieve stress around compact defects by plastic deformation. The BFI is determined[73,74] by comparing the tensile strength of a compact, σ_T, with that of a compact with a small hole (stress concentrator) in it, σ_{To}, using the tensile test described above. A hole in the center of a compact weakens it. If a material is very brittle, theoretical considerations show that the tensile strength of a tablet with a hole in it will be about one-third that of a "defect free" tablet. If, however, the material can relieve stress, then the strength of the compact with a hole in it will approach that of a compact with no hole. The BFI is defined such that very brittle compacts have a *BFI* of 1, and very nonbrittle materials have a *BFI* close to 0; it is calculated in Eq. (10.10).[73] *BFI* values less than 0.3 (range 0−1) are indicative of relatively nonbrittle materials.

$$BFI = 0.5 \cdot \left[\frac{\sigma_T}{\sigma_{To}} - 1\right] \tag{10.10}$$

10.4.3.8 Viscoelastic index

Hiestand and coworkers have further refined the concept of BI to include both a worst-case and a best-case BI.[75] The BI is determined under different experimental conditions: the rate at which the permanent dent is made in a compact is varied such that the viscoelastic properties of the material are assessed. If a material is very viscoelastic, there is substantial stress relaxation with time. It is reasonable to expect, then, that tablets that are slowly deformed during the determination of the hardness, H, may retain more of the bonded area than tablets that are rapidly deformed (ie, as in the PID). This is because some of the stresses developed during compaction will have a chance to be relieved. The dynamic bonding index (BI_d), sometimes called the worst case BI, is determined using a the PID for measuring the indentation hardness (H_d). On the other hand, the quasi-static bonding index (BI_{qs}), also sometimes referred to as the best case BI, is measured using a "quasi-static" or slow method for measuring indentation hardness (H_{qs}). The dynamic and quasi-static BI is calculated as previously described. The viscoelastic index (VE) is defined as the ratio of the dynamic to quasi-static indentation hardness:

$$VE = \frac{H_d}{H_{qs}} = \frac{BI_{qs}}{BI_d} \tag{10.11}$$

10.4.3.9 Application of Quasi-static testing to formulation development

The application of quasi-static testing methods and interpretation has been discussed extensively in the scientific literature. In addition to the pioneering work of Hiestand and coworkers,[60–62,73–77] additional research discussing the application of this methodology is available.[33,63,80–89] Benefits of a complete characterization of the mechanical properties of both the active ingredient and the excipients used in the formulation include:

- fundamental understanding of critical mechanical properties of the active ingredient and excipients;
- identification of mechanical property deficiencies and attributes;
- selection of excipients that can overcome deficiencies of active ingredient;
- identification of lot-to-lot variations in materials;
- identification of potential manufacturing problems associated with tableting process (Retrieved from http://www.authorstream.com/Presentation/naveen.gokanapudi-1764246-preformulatio).

The reader is directed to the literature for a thorough discussion of the application of mechanical property characterization to formulation development. A fundamental understanding of the mechanical properties is essential to understanding compaction properties and the tableting process.[60,61,75–77,79] Development of mathematical models of mixtures has been used by Amidon,[90,91] and others,[86,87,92] to identify the type and quantity of excipient required to produce tablet formulations that have acceptable manufacturing properties. Fig. 10.11 shows mechanical properties of binary mixtures of microcrystalline cellulose and lactose spray process.[90,91,93] From this figure, one can see that the mechanical properties of a mixture may be estimated knowing the mechanical

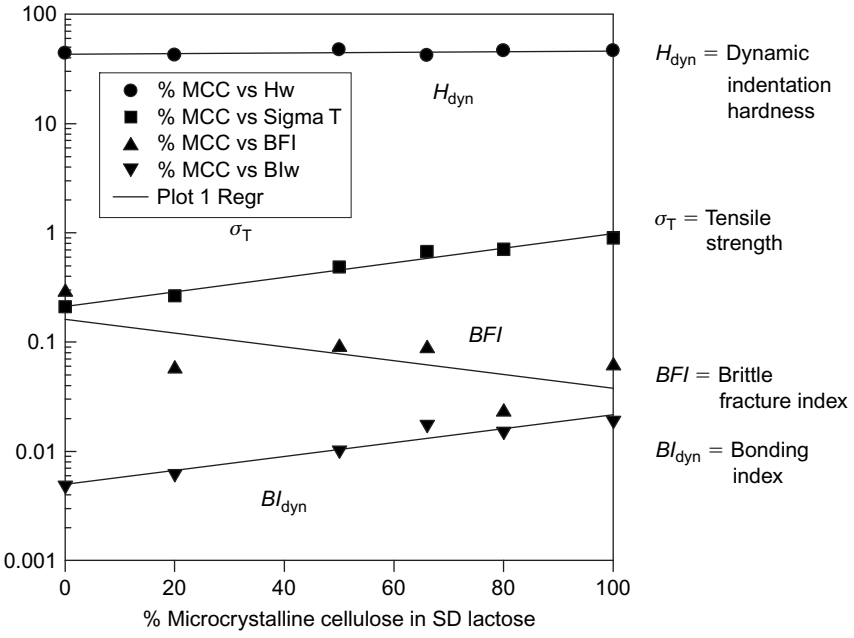


FIGURE 10.11 Mechanical properties of mixtures of microcrystalline cellulose and lactose spray process.

properties of the two individual components. While the mechanical properties of mixtures are complicated,[94,95] estimating the properties of mixtures has been successfully used to identify suitable excipient types and quantities. A simplified equation for binary mixtures is given in Eq. (10.12) or Eq. (10.13):

$$\Phi_{\text{mixture}} = x(\Phi_A) + (1-x)(\Phi_B) \tag{10.12}$$

$$\log(\Phi_{\text{mixture}}) = x\log(\Phi_A) + (1-x)\log(\Phi_B) \tag{10.13}$$

where:

Φ_{mixture} is the mechanical property of interest for the mixture at the reference SF
Φ_A, Φ_B are the properties of the pure components of the binary mixture at the same reference SF
x is the fraction of component A.

While either equation may be used, experience has shown that the log-linear relationship shown in Eq. (10.13) often produces better predictions for mixtures of common lactose-microcrystalline cellulose-based formulations. Further work is needed, however, to explore the science and predictability of mixtures.

For ternary mixtures, that is, for mixtures containing an API as one component and a second "placebo" component containing two excipients, it is possible to estimate the mechanical properties of a mixture by extending this concept, as shown in Eq. (10.14).

$$\log(\Phi_{\text{formulation}}) = y\log(\Phi_{\text{API}}) + (1-y) \cdot([x]\log(\Phi_A) + [1-x]\log\Phi_B) \tag{10.14}$$

where:

y is the fraction of API in the blend
$(1-y)$ is the fraction of the blend that constitutes the placebo component.

While a simplification, this approach has been used to predict the properties of mixtures.[90,91,93] An example of the predicted mechanical properties of a ternary blend of API, and a placebo component consisting of microcrystalline cellulose and lactose spray process, is shown in Fig. 10.12. A sound understanding of the mechanical properties of the individual components, and the range of desirable mechanical property values, allows for a rational selection of excipient types, grades and quantities.

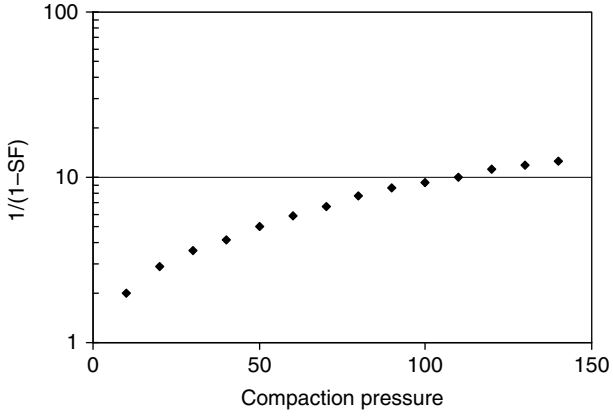


FIGURE 10.12 Heckel plot.

10.4.4 Dynamic testing

The most commonly used methods of studying the mechanical properties of solids under dynamic conditions include the use of the following instrumented equipment:

- Hydraulic press;
- Eccentric (single station) tablet press;
- Rotary tablet press;
- Compaction simulator;
- Compaction emulator (eg, Presster).

Accurate measurements of force and distance require careful construction of equipment and placement of instrumentation; several publications are available discussing these aspects.[96,97] Appropriate care is also needed during calibration, and for correction of measurements due to machine dimensional changes; tablet punches, for example, deform elastically, and these changes should be considered.

While any of these instrumented presses may be used, the most commonly used for dynamic testing is the compaction simulator. The compaction simulator,[98] because of its sophisticated control and monitoring of the compaction process, offers the greatest flexibility in compression conditions. It is possible, for example, to carry out compression under constant velocity conditions for the compression and decompression phases (ie, saw-tooth compression profile)—something not possible with other presses. Also of particular note is the demonstrated utility of a "linear" tablet press emulator (Presster), that offers many of the advantages of a conventional compaction simulator in ease of use and experimental flexibility, as well as its ability to simulate virtually any commercially available rotary tablet press.[99] One limitation of the Presster, though it may also be viewed as a benefit, is that it uses compression rolls of the same dimensions as those on a rotary tablet press. Therefore, compression profiles are limited to those described by conventional tablet press geometries of the tablet punch tooling as it moves under the compression rolls.

With proper instrumentation, "in die" or "at pressure" measurements may be made during the compaction process. For example, the compaction pressure versus tablet porosity (compressibility) may be determined for a single tablet. This information, obtained under

these dynamic testing conditions, can be used to generate an "in die" Heckel plot. Alternatively, measurements may be made on compacts after the tablet is removed from the die. These are "out of die" or "at zero pressure" measurements. Both in die and out of die methods have their advantages. Among the advantages of out of die measurements is that they represent the "final product" after decompression and ejection. The out of die properties are those by which the dosage form are ultimately judged by the formulator, by quality assurance, and by the consumer.

A variety of pressure–porosity equations have been derived over the years.[100–102] The most commonly referenced of them is that of Heckel.[102] The Heckel equation (sometimes referred to as the Athy–Heckel equation) was derived assuming that the change in SF (ie, relative density) with compaction pressure is proportional to the porosity of the compact. Therefore, as porosity approaches zero, the change in SF with compaction pressure, $d\rho/dP$, approaches zero. Therefore:

$$\frac{d\rho_r}{dP} \sim \varepsilon = (1-\rho_r) \tag{10.15}$$

And, integrating Eq. (10.15), results in the classic Heckel equation:

$$\ln\left(\frac{1}{1-\rho_r}\right) = kP + A \tag{10.16}$$

where k is a measure of the plasticity of the material. It is related to the yield strength, Y, of a material by Eq. (10.17).[102,103]

$$k = \frac{1}{3y} \tag{10.17}$$

The constants in the Heckel plot are typically determined by regression analysis of the terminal linear portion of a plot of $\ln(1/(1-\rho_r))$ versus P (see Fig. 10.13). The yield strength, Y, of the material under the dynamic conditions of the test is a measure of the deformability. In addition to yield strength, the shape of the Heckel plot has been used to distinguish volume reduction mechanisms.[104] Three types or families of curves are considered to reflect materials that undergo consolidation primarily by: (1) plastic deformation, (2) fragmentation, or (3) a variation of (1) which is plastic flow with no initial particle rearrangement.

Additional information regarding the compaction process may be obtained using dynamic testing conditions, including work of compaction, work recovered during decompression, work to overcome die wall friction, etc. A detailed discussion of these opportunities is beyond the scope of this chapter, and the reader is directed to the literature for further information. While very valuable as a research tool, the quantitative use of pressure–porosity measurements and analysis beyond the determination of yield pressure does not appear to be used routinely during formulation development and optimization.

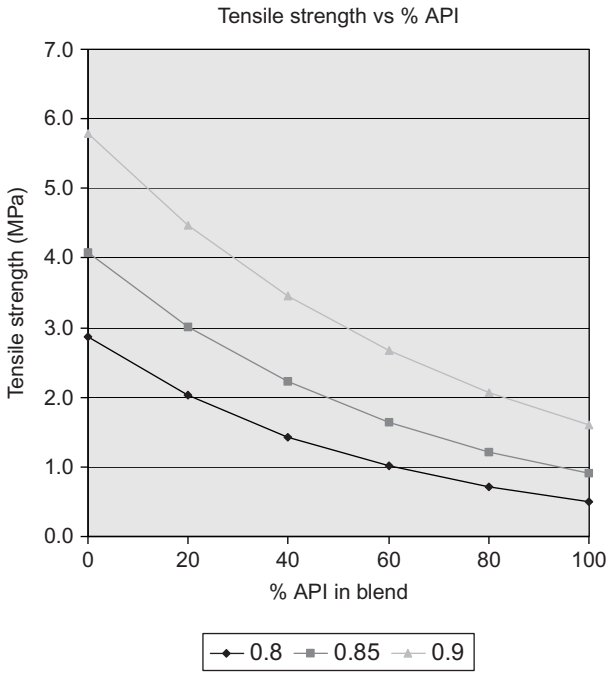


FIGURE 10.13 Predicted mechanical properties of a ternary blend of API, micro crystalline cellulose and lactose spray process.

10.4.4.1 Application of dynamic testing to formulation development

There are a number of reports of the use of dynamic testing of active ingredients and excipients in the literature. There are two key benefits of dynamic testing: (1) the properties can be determined under dynamic conditions representing those in a production environment and (2) small quantities are typically required (2–10 g). In contrast, disadvantages include the difficulty of factoring out the individual mechanical property "components" that, combined, determine how a material behaves during compaction.

The relationships between compaction pressure, tensile strength, and SF are critical to understanding and characterizing the compaction process. The relationship between these three parameters is described as:

- Compactibility: relationship between tensile strength and SF;
- Tabletability: relationship between tensile strength and compression pressure;
- Compressibility: relationship between compaction pressure and SF or porosity;
- Manufacturability: relationship between tablet crushing force and compression force.

Representative compactibility, tabletability, compressibility, and manufacturability profiles for a compactable excipient are shown in Figs. 10.14–10.17. The compactibility, tabletability, and compressibility profiles form the three faces of a three-dimensional plot as shown in Fig. 10.18.[63] The Presster compaction emulator was used for these studies, although other properly instrumented presses can also be used. The compaction emulator was set up to emulate a Killian RST tablet machine (250 mm compression rolls) with a 27 ms dwell time (corresponding to 28,800 tablets/hour) using 10 mm diameter flat-faced round punches with no precompression force.

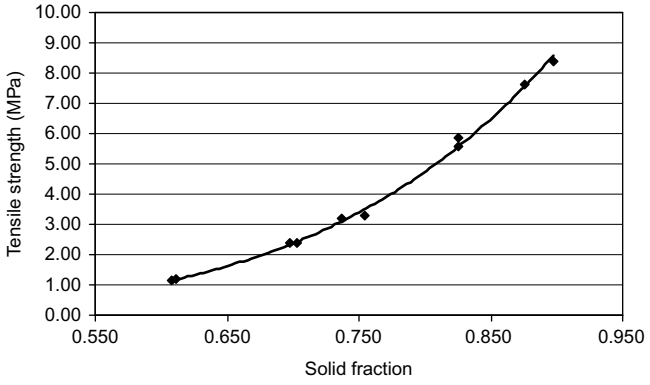


FIGURE 10.14 Compactibility profile using a compaction emulator.

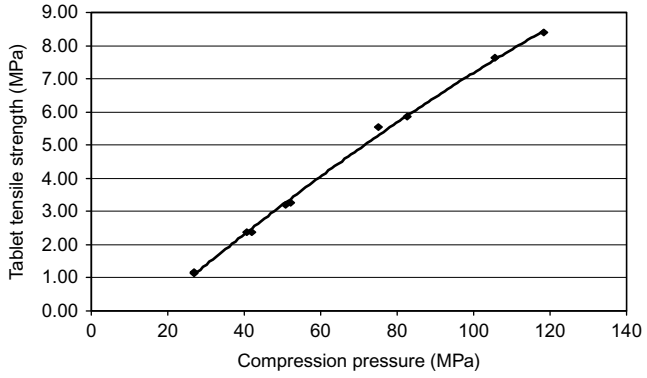


FIGURE 10.15 Tabletability profile using a compaction emulator.

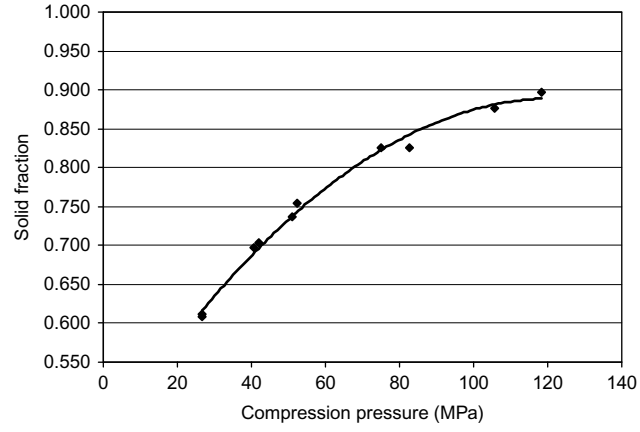


FIGURE 10.16 Compressibility profile using a compaction emulator.

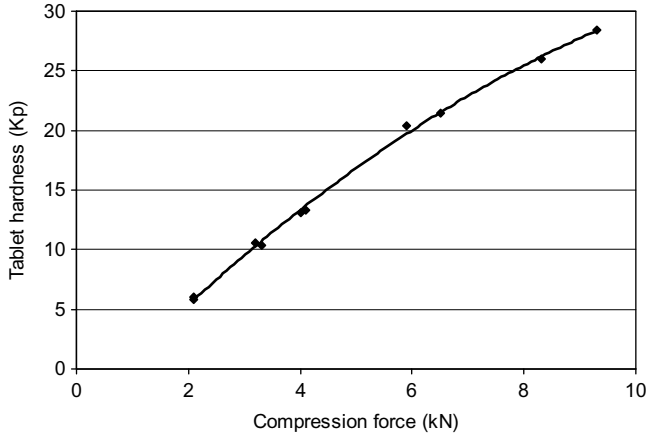


FIGURE 10.17 Manufacturability profile using a compaction emulator.

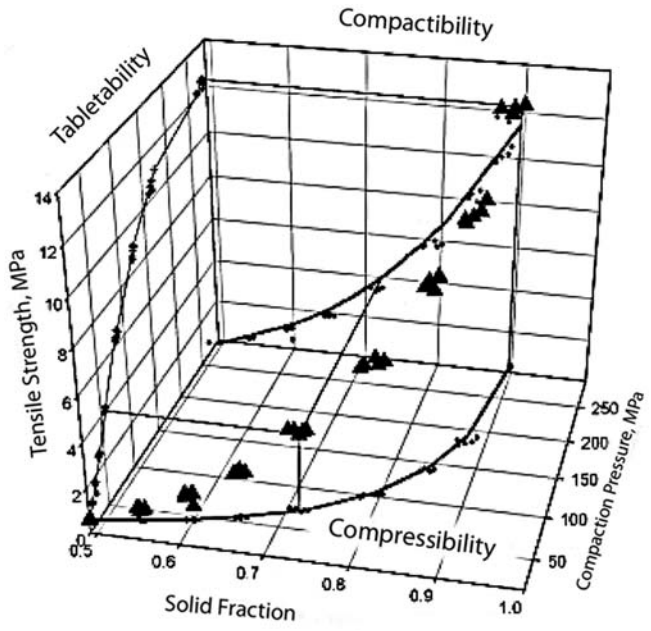


FIGURE 10.18 Three-dimensional tablet tensile strength, solid fraction, and compaction pressure curve.

Compactibility is the ability of a powder to be transformed into tablets with a resulting strength.[105] It is represented by a plot of tensile strength versus SF. The compactibility is the most valuable of the three properties since it reflects the two most important effects of applied pressure: tablet strength, and SF. A representative compactibility profile of an excipient is shown in

Fig. 10.14. If one can achieve an acceptable tensile strength at an acceptable SF with the application of pressure, a satisfactory tablet can be produced. Compactibility plots are largely independent of the process by which compacts are made since only measured tablet properties (tensile strength and SF) are involved. Compactibility plots are useful as a tool to compare formulations made on different equipment. If the formulations are the "same," then the "same" compactibility plots will be obtained.[63]

Tabletability is the capacity of a powder to be transformed into a tablet of specified strength under the effect of compaction pressure.[105] It is represented by a plot of tensile strength versus compaction pressure. Tabletability describes the effectiveness of the applied pressure in increasing the tensile strength of the tablet, and demonstrates the relationship between the cause (the compaction pressure), and the effect (the strength of the compact) (see Fig. 10.15). Normally, a higher compaction pressure makes a stronger tablet. However, this relationship is often found to be speed dependent. Also, at high pressures, some materials may have lower tensile strength due to overcompaction.[106] Characterization of the tabletability provides excellent insight into the compaction process and mechanical properties of a material.

Compressibility is the ability of a material to undergo a reduction in volume as a result of an applied pressure.[105,107] It is a measure of the ease with which a powder bed undergoes volume reduction under compaction pressure; it is represented by a plot showing the reduction of tablet porosity (ie, the increase in SF) with increasing compaction pressure (Fig. 10.16). Compressibility is often described by the Heckel equation.[102] Heckel plots, for example, have been widely used to assess the mechanism of deformation, and as a tool to estimate yield pressure. It is also well-known that tablet porosity is an important parameter, for example, in tablet disintegration and dissolution, since some porosity is often necessary to facilitate liquid penetration into tablets.[103,108]

Manufacturability, a plot closely related to tabletability, shows the relationship between the tablet crushing force (related to tensile strength), and compaction force (related to compression pressure). The manufacturability profile (Fig. 10.17) is commonly considered by formulation scientists since it reflects the "measured" properties of a dosage form during manufacturing (tablet crushing strength and compression force). In general, however, pressure and tensile strength are preferred parameters to consider.

In summary, characterization of the compactibility, tabletability, compressibility, and manufacturability of a formulation provides valuable information of the compaction process and the prospects for a successful tableting process in manufacturing. Obtaining tablets with adequate tensile strength at a reasonable SF with acceptable compression pressure is the key to success. Robust formulations must not be on the "edge;" that is, they should provide the manufacturing scientist with the ability to adjust compression pressure to achieve the desired tensile strength and still maintain the SF in a desirable range, such that the tablet performs as required.

10.5 CONCLUSIONS

As timelines become tighter and shortened, it has become more important than ever to quickly and efficiently characterize the critical properties of materials that will influence product development and performance. In this chapter, a discussion of those particle, powder, and compact properties that are most important in developing solid dosage forms has been discussed. The focus has been on methods that yield important information, yet require small quantities of materials. With a sound understanding of these properties, formulation development can proceed most efficiently and scientifically with greater success. Tomorrow's formulation and process scientists will require a sound understanding of these pharmaceutical material science principles, and must be able to apply them to the design and development of dosage forms in an efficient and scientifically rigorous way. The beauty of science is the knowledge and ability it gives us to predict the future reliably. As formulation and process scientists, the future we need to predict accurately is that of a consistent, reliable, manufacturable product that performs as expected.

References

1. Brittain HG. Evaluation of the particle size distribution of pharmaceutical solids. In: Brittain HG, editor. *Profiles of drug substances, excipients, and related methodology*, vol. 31. Amsterdam: Elsevier Academic Press; 2004. p. 379–419.
2. Allen T. 5th ed. *Particle size measurement*, vol. 1. New York, NY: Chapman & Hall; 1997.
3. Stanforth J. Particle-size analysis. In: Aulton M, editor. *Pharmaceutics the science of dosage form design*. New York, NY: Chapman & Hall; 2004. p. 152–65.
4. US Pharmacopeia 27. *US pharmacopeial convention*. Rockville, MD; 2006. p. 2716–2717.
5. Newman AW, Brittain HG. Particle morphology: optical and electron microscopies. In: Brittain HG, editor. *Physical characterization of pharmaceutical solids*. New York, NY: Marcel Dekker; 1995. p. 127–56.
6. Operation and Maintenance Manual. *Gilsonic siever model GA-8*. Lewis Center, OH: Gilson Company, Inc; 2004.
7. US Pharmacopeia 27. US pharmacopeial convention. Rockville, MD; 2006. p. 2720–2722.

8. Rohrs BR, Amidon GE, Meury RH, Secreast PJ, King HM, Skoug CJ. Particle size limits to meet USP content uniformity criteria for tablets and capsules. *J Pharm Sci* 2006;**95**(5):1049–59.
9. Dressman JB, Fleisher D. Mixing tank model for predicting dissolution rate control of oral absorption. *J Pharm Sci* 1986;**75**:109–16.
10. Higuchi WI, Hiestand EN. Dissolution rates of finely divided drug powders I: effect of a distribution of particle sizes in a diffusion-controlled process. *J Pharm Sci* 1963;**52**:167–71.
11. Hintz RJ, Johnson KC. The effect of particle size distribution on dissolution rate and oral absorption. *Int J Pharm* 1988;**51**:9–17.
12. Rohrs BR, Amidon GE. Particle engineering: a formulator's perspective. In: *AAPS Arden House Conference*. New York: Harriman; 2005.
13. US Pharmacopeia 27. US pharmacopeial convention. Rockville, MD; 2006. p. 2669–2670.
14. Mohanmmadi MS, Harnby N. Bulk density modeling as a means of typifying the microstructure and flow characteristics of cohesive powders. *Powder Technol* 1997;**92**:1–8.
15. Abdullah EC, Geldart D. The use of bulk density measurements as flowability indicators. *Powder Technol* 1999;**102**:151–65.
16. US Pharmacopeia 27. US pharmacopeial convention. Rockville, MD; 2006. p. 2638–2639.
17. Suzuki M, Sato H, Hasegawa M, Hirota M. Effect of size distribution on tapping properties of fine powder. *Powder Technol* 2001;**118**:53–7.
18. Amidon GE, Houghton ME. Powder flow testing in preformulation and formulation development. *Pharm Manuf* 1985;**2**(7):20–31.
19. Köhler T. Influence of particle size distribution on the flow behaviour of fine powders. *Part Part Syst Charact* 1991;**8**:101–4.
20. Shinohara K. Effect of particle shape on angle of internal friction by triaxial compression test. *Powder Technol* 2000;**107**:131–6.
21. Podczeck F. The influence of particle size and shape on the angle of internal friction and the flow factor of unlubricated and lubricated powders. *Int J Pharm* 1996;**144**:187–94.
22. Kachrimanis K, Karamyan V, Malamataris S. Artificial neural networks (ANNs) and modeling of powder flow. *Int J Pharm* 2003;**250**:12–23.
23. Moreno-Atanasio R, Antony SJ, Ghadiri M. Analysis of flowability of cohesive powders using Distinct Element Method. *Powder Technol* 2005;**158**:51–7.
24. Tomas RJ. Assessment of mechanical properties of cohesive particulate solids. Part 2: powder flow criteria. *Part Sci Technol* 2001;**19**:111–29.
25. Carr RL. Classifying flow properties of solids. *Chem Eng* 1965;**72**:69–70.
26. Podczeck F, Newton JM. Powder filling into hard gelatine capsules on a tamp filling machine. *Int J Pharm* 1999;**185**:237–54.
27. Grey RO, Beddow JK. On the Hausner Ratio and its relationship to some properties of metal powders. *Powder Technol* 1969;**2**(6):323–6.
28. Ho R, Bagster DF, Crooks MJ. Flow studies on directly compressible tablet vehicles. *Drug Develop Indust Pharm* 1977;**3**:475.
29. Nelson E. Measurement of the repose angle of a tablet granulation. *J Am Pharm Assoc Sci Ed* 1955;**44**:435.
30. Armstrong NA, Griffiths RV. The effects of moisture on the flow properties and compression of phenacetin paracetomol and dextrose monohydrate. *Pharm Acta Helv* 1970;**45**:692.
31. US Pharmacopeia 27. US pharmacopeial convention. Rockville, MD; 2006. p. 3017–3020.
32. Khanam J, Nanda A. Flow of granules through cylindrical hopper. *Powder Technol* 2005;**150**:30–5.
33. Amidon GE. Physical and mechanical property characterization of powders. In: Brittain HG, editor. *Physical characterization of pharmaceutical solids*, Vol. 70. New York, NY: Dekker; 1995. p. 281–319.
34. Grossmann J, Tomas J. Flow properties of cohesive powders tested by a press shear cell. *Part Sci Technol* 2006;**24**:353–67.
35. Schwedes J, Schulze D. Measurement of flow properties of bulk solids. *Powder Technol* 1990;**61**:59–68.
36. Hiestand EN, Wilcox CJ. Some measurements of friction in simple powder beds. *J Pharm Sci* 1968;**57**:1421–7.
37. Marchall K, Sixsmith D. The flow properties of microcrystalline cellulose powders. *J Pharm Pharmacol* 1976;**28**:770.
38. Fizpatrick JJ, Barringer SA, Iqbal T. Flow property measurement of food powders and sensitivity of Jenike's hopper design methodology to the measured values. *J Food Eng* 2004;**61**:399–405.
39. USP39-NF34, <1063> Shear cell methodology for powder flow testing. *Pharm Forum* 2015;**41**(5):8348–59.
40. Schulze D. The behavior of powders and bulk solids. Fundamentals of Bulk Solid Mechanics. <www.dietmar-schulze.de/grdle1.html>.
41. Podczeck F, Mia Y. The influence of particle size and shape on the angle of internal friction and the flow factor of unlubricated and lubricated powders. *Int J Pharm* 1996 29 November;**144**(2):187–94.
42. Podczeck F, Newton JM. Powder and capsule filling properties of lubricated granulated cellulose powder. *Eur J Pharm Biopharm* 2000;**50**:373–7.
43. Nyqvist H. Prediction of weight variation in tablet production from shear cell measurements. *Acta Pharm Suec* 1982;**19**:413–20.
44. Fitzpatrick JJ, Barringer SA, Iqbal T. Flow property measurement of food powders and sensitivity of Jenike's hopper design methodology to the measured values. *J Food Eng* 2004;**61**:399–405.
45. Schulze D. A new ring shear tester for flowability and time consolidation measurements. In: Proc. 1st international particle technology forum, August. Denver/CO; 1994. p. 11–16.
46. Jenike AW. Storage and flow of solids. Bulletin 123. Engineering Experiment Station, University of Utah; 1964.
47. Dietmar Schulze. Ring Shear Tester RST-XS Operating instructions; 2002–2004.
48. ASTM International. Standard shear test method for bulk solids using the Schulze ring shear tester, D 6773–02.
49. Schwedes J. Testers for measuring flow properties of particulate solids. *Powder Handl Process* 2000;**12**(4):337–54.
50. ASTM International. Standard test method for shear testing of bulk solids using the Jenike shear cell, D 6128–00.
51. Hiestand EN, Wells JE. A simplified shear cell apparatus and procedure. In: Proceedings of the International Powder and Bulk Solids Handling and Process Conference. Rosemont, IL; 1977 May. p. 244.
52. Kaye BH. Characterizing the flowability of a powder using the concepts of fractal geometry and chaos theory. *Part Part Syst Charact* 1997;**14**:53–66.
53. Hancock BC. Development of a robust procedure for assessing powder flow using a commercial avalanche testing instrument. *J Pharm Biomed Anal* 2004;**35**:979–90.
54. Hickey AJ, Concessio NM. Flow properties of selected pharmaceutical powders from a vibrating spatula. *Part Part Syst Charact* 1994;**11**:457–62.
55. Rastogi S, Klinzing GE. *Part Part Syst Charact* 1994;**11**:453–6.
56. Doherty R, Sahajpal H, Poynter R, Lee Y. *J Pharm Pharmacol* 1999;**51**:S323.
57. Boothroyd EM, Doherty RA, Poynter R, Ticehurst MD. Comparison of blend flow measured on the Aero-Flow™ with tablet weight uniformity. *J Pharm Pharmacol* 2000;**52**:174.
58. Boothroyd EM, Doherty RA, Poynter R, Ticehurst M. *J Pharm Pharmacol* 2000;**52**:174S.
59. Bhattachar SN, Hedden DB, Olsofsky AM, Qu X, Hsieh WY, Canter KG. Evaluation of the vibratory feeder method for assessment of powder flow properties. *Int J Pharm* 2004;**269**:385–92.

60. Hiestand EN. *Mechanics and physical principles for powders and compacts*. West Lafayette, IN: SSCI Inc; 2000.
61. Hiestand EN. Rationale for and the measurement of tableting indices. In: Alderborn G, Nystrom C, editors. *Pharmaceutical powder compaction technology*, vol. 71. New York, NY: Marcel Dekker, Inc; 1996. p. 219–44.
62. Hiestand EN, Bane JM, et al. Impact test for hardness of compressed powder compacts. *J Pharm Sci* 1971;**60**(5):758–63.
63. Tye CK, Sun C, Amidon GE. Evaluation of the effects of tableting speed on the relationships between compaction pressure, tablet tensile strength, and tablet solid fraction. *J Pharm Sci* 2005;**94**(3):465–72.
64. Rees JE, Rue PJ. Time-dependent deformation of some direct compression excipients. *J Pharm Pharmacol* 1978;**30**:601–7.
65. David ST, Augsburger LL. Plastic flow during compression of directly compressible fillers and its effect on tablet strength. *J Pharm Sci* 1977;**66**(2):155–9.
66. Roberts RJ, Rowe RC. The effect of the relationship between punch velocity and particle size on the compaction behaviour of materials with varying deformation mechanisms. *J Pharm Pharmacol* 1986;**38**:567–71.
67. Ishino R, Yoshino H, Hirakawa Y, Noda K. Influence of tabletting speed on compactibility and compressibility of two direct compressible powders under high speed compression. *Chem Pharm Bull* 1990;**38**(7):1987–92.
68. Ruegger CE, Celik M. The effect of compression and decompression speed on the mechanical strength of compacts. *Pharm Develop Technol* 2000;**5**(4):485–94.
69. Armstrong NA, Palfrey LP. The effect of machine speed on the consolidation of four directly compressible tablet diluents. *J Pharm Pharmacol* 1989;**41**:149–51.
70. Hancock BC, Colvin JT, Mullarney MP, Zinchuk AV. The relative densities of pharmaceutical powders, blends, dry granulations, and immediate-release tablets. *Pharm Technol* 2003;**27**(4):64–80.
71. Maarschalk KV, Vromans H, Groenendijk W, Bolhuis GK, Lerk CF. Effect of water on deformation and bonding of pregelatinized starch compacts. *Eur J Pharm Biopharm* 1997;**44**(3):253–60.
72. Alderborn G, Nystrom C, editors. *Pharmaceutical powder compaction technology*, vol. 71. New York, NY: Marcel Dekker, Inc; 1996.
73. Hiestand E, Wells JE, Peot CB, Ochs JF. Physical processes of tableting. *J Pharm Sci* 1977;**66**(4):510–19.
74. Hiestand EN, Bane JM, Strzelinski EP. Impact test for hardness of compressed powder compacts. *J Pharm Sci* 1971;**60**(5):758–63.
75. Hiestand EN, Smith DP. Indices of tableting performance. *Powder Technol* 1984;**38**:145–59.
76. Hiestand EN, Smith DP. Tablet Bond I. A theoretical model. *Int J Pharm* 1991;**67**:217–29.
77. Hiestand EN, Smith DP. Tablet Bond II. Experimental check of model. *Int J Pharm* 1991;**67**:231–46.
78. Fell JT, Newton JM. Determination of tablet strength by the diametral-compression test. *J Pharm Sci* 1970;**59**(5):688–91.
79. He X, Secreast PJ, Amidon GE. Mechanistic study of the effect of roller compaction and lubricant on tablet mechanical strength. *J Pharm Sci* 2007;**96**(5):1342–55.
80. Mullarney MP, et al. The powder flow and compact mechanical properties of sucrose and three high-intensity sweeteners used in chewable tablets. *Int J Pharm* 2003;**257**(12):227–36.
81. Boldyreva EV, Dmitriev V, Hancock BC. Effect of pressure up to 5.5 GPa on dry powder samples of chlorpropamide form-A. *Int J Pharm* 2006;**327**(1–2):51–7.
82. Hancock BC, et al. The powder flow and compact mechanical properties of two recently developed matrix-forming polymers. *J Pharm Pharmacol* 2001;**53**(9):1193–9.
83. Hancock BC, Dalton CR, Clas S. Micro-scale measurement of the mechanical properties of compressed pharmaceutical powders 2: the dynamic moduli of microcrystalline cellulose. *Int J Pharm* 2001;**228**(1–2):139–45.
84. Wu CY, Best SM, Bentham AC, Hancock BC, Bonfield W. Predicting the tensile strength of compacted multi-component mixtures of pharmaceutical powders. *Pharm Res* 2006;**23**(8):1898–905.
85. Mullarney MP, Hancock BC, Narajan P. Improving the prediction of exceptionally poor tableting performance: an investigation of Hiestand's "special case". *J Pharm Sci* 2004;**93**(8):2017–21.
86. Wu CY, Best SM, Bentham AC, Hancock BC, Bonfield W. A simple predictive model for the tensile strength of binary tablets. *Eur J Pharm Sci* 2005;**25**(2–3):331–6.
87. Wu CY, Ruddy OM, Bentham AC, Hancock BC, Best SM, Elliott JA. Modelling the mechanical behaviour of pharmaceutical powders during compaction. *Powder Technol* 2005;**152**(1–3):107–17.
88. Hancock BC, Carlson GT, Ladipo DD, Langdon BA, Mullarney MP. Comparison of the mechanical properties of the crystalline and amorphous forms of a drug substance. *Int J Pharm* 2002;**241**(1):73–85.
89. Amidon GE, Houghton ME. The effect of moisture on the mechanical and powder flow properties of microcrystalline cellulose. *Pharm Res* 1995;**12**(6):923–9.
90. Amidon GE. Data driven formulation development using material sparing methods. In: Garnet Peck Symposium, Purdue University, West Lafayette, IN; 2006 September 20–21.
91. Amidon GE. Mechanical property characterization of API, excipients and formulations during development. In: AAPS Pharmaceutics and Drug Delivery Conference, April, Washington, DC; 2002.
92. Leuenberger H, Leu R. Formation of a tablet: a site and bond percolation phenomenon. *J Pharm Sci* 1992;**81**(10):976–82.
93. Amidon GE. Mechanical property characterization of pharmaceutical mixtures. Department of Materials Science, University of Cambridge, UK; 2006 August 2.
94. Fell JT. Compaction properties of binary mixtures. In: Alderborn G, Nystrom C, editors. *Pharmaceutical powder compaction technology*, vol. 71. New York, NY: Marcel Dekker, Inc; 1996. p. 501–16.
95. Wurster DE, Majuru S, Oh E. Prediction of the Hiestand bonding indices of binary powder mixtures from single-component bonding indices. *Pharm Develop Technol* 1999;**49**(1):65–70.
96. Lammens RF, Fedin VP. Modern press instrumentation. *J Med Pharm Chem* 1978;**12**:263.
97. Paronen P, Ilkka J. Porosity-pressure functions. In: Alderborn G, Nystrom C, editors. *Pharmaceutical powder compaction technology*, vol. 71. New York, NY: Marcel Dekker, Inc; 1996.
98. Huxley Bertram Engineering Ltd. Cambridge, UK. <http://www.powder-compaction.com/compaction-simulators.htm>
99. Metropolitan Computing Corp. <http://www.mcc-online.com/presster.htm>.
100. Cooper AR, Eaton LE. Compaction behavior of several ceramic powders. *J Am Ceram Soc* 1962;**45**:97–101.
101. Nelson E, Busse LW, Higuchi T. The physics of tablet compression VII Determination of energy expenditure in the tablet compression process. *J Am Pharm Assoc Sci Ed* 1955;**44**:223.
102. Heckel RW. Density–pressure relationships in powder compaction. *Trans Metall Soc AIME* 1961;**221**:671–5.

103. Hersey JA, Rees J. Deformation of particles during briquetting. *Nature (PhysSci)* 1971;**230**:96.
104. York P, Pilpel N. The tensile strength and compression behavior of lactose, four fatty acids and their mixtures in relation to tableting. *J Pharm Pharmacol* 1973;**25**:1.
105. Joiris E, Di Martino P, Berneron C, Guyot-Hermann AM, Guyot JC. Compression behavior of orthorhombic paracetamol. *Pharm Res* 1998;**15**(7):1122–30.
106. Maarschalk KV, Zuurman K, Vromans H, Bolhuis GK, Lerk CF. Porosity expansion of tablets as a result of bonding and deformation of particulate solids. *Int J Pharm* 1996;**140**(2):185–93.
107. Duberg M, Nystrom C. Studies on direct compression of tablets. VI. Evaluation of methods for the estimation of particle fragmentation during compression. *Acta Pharm Suec* 1982;**19**:421–36.
108. Celik M, Ong JTH, Chowhan ZT, Samuel GJ. Compaction simulator studies of a new drug substance: effect of particle size and shape, and its binary mixtures with microcrystalline cellulose. *Pharm Develop Technol* 1996;**1**(2):119–26.

This page intentionally left blank

PART II

BIOPHARMACEUTICAL AND PHARMACOKINETIC EVALUATIONS OF DRUG MOLECULES AND DOSAGE FORMS

This page intentionally left blank

CHAPTER

11

Oral Absorption Basics: Pathways and Physicochemical and Biological Factors Affecting Absorption

L. Zhu[1,4], *L. Lu*[1], *S. Wang*[3], *J. Wu*[1], *J. Shi*[1], *T. Yan*[1], *C. Xie*[1], *Q. Li*[1], *M. Hu*[1,2] *and Z. Liu*[1]

[1]Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, PR China
[2]University of Houston, Houston, TX, United States [3]Worldwide Research and Development, Pfizer, Cambridge, MA, United States [4]AbbVie Inc., North Chicago, MA, United States

11.1 BARRIERS TO ORAL DRUG DELIVERY

Oral drug delivery is the most popular and economical route of drug administration. Barriers to oral drug delivery include solubility and dissolution, absorption, presystemic metabolism, and excretion. For drugs whose targets are easily accessible, the portion of drugs reaching systemic circulation after overcoming all of these barriers is considered bioavailable. Certain drugs that must reach hard-to-access target organs, such as the central nervous system (CNS) and drug-resistant cancer cells; and must be transported across additional barriers (eg, the blood–brain barrier) to their targets. Consequently, these barriers may represent an additional hurdle that should be overcome for delivering orally administered drugs.

All oral drugs must generally be dissolved in the aqueous environment of the gastrointestinal (GI) tract before they can be absorbed. This seemingly simple process is actually a significant challenge to the development of new oral drugs, especially lipophilic molecules that are poorly soluble in aqueous solutions. Several chapters in this book deal with this issue; thus, we will not discuss them in detail here. An important issue to recognize is that the transit time for solid-dosage forms or undissolved drug solids in the upper GI tract (ie, stomach and small intestine) typically ranges from 2 to 6 hours.[1] Therefore, dissolution rates must be sufficiently high to dissolve these solids for absorption.

Once dissolved in the gastric fluid, drug molecules must stay soluble throughout the GI tract. This requirement is not trivial because the pH of the GI tract changes from highly acidic (pH1.5) in the stomach to slightly basic (pH8) in the lower small intestine and colon; as this change could alter the solubility of numerous drugs, this factor should be considered when designing solubility and dissolution experiments.

Dissolved drugs are absorbed via different transcellular and paracellular pathways across the intestinal tract. A paracellular pathway is defined as drug transport between cells, whereas a transcellular pathway is defined as drug transport across the cells (through apical and basolateral membranes). Transcellular pathways include passive diffusion and carrier-mediated transport, whereas paracellular pathways mainly involve passive diffusion.

Drugs are primarily absorbed in the small intestine, where the absorption area is very large (100 m^2 or an area equivalent to that of a tennis court). Absorption in the colon is generally limited and can be highly variable because of undigested materials, bacteria, and variable water content. Given these factors, GI transit plays a significant role in the absorption of drugs.

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00011-X

After absorption into intestinal cells, some drugs may be subjected to first-pass metabolism by various intestinal metabolic enzymes. Escaping the intestinal metabolism does not guarantee bioavailability because the drugs absorbed from the stomach and intestine (large and small), with the exception of those absorbed at the terminal rectum, are subject to additional first-pass metabolism by the liver. Metabolites formed in the intestinal cells can be directly or indirectly excreted back to the lumen after they are taken up by the liver cells and excreted via bile. Metabolites formed in hepatocytes can also be excreted via bile. Alternatively, all metabolites formed in enterocytes or hepatocytes may stay in the blood, where they are eliminated via the kidneys.

11.1.1 Intestinal barrier

The intestinal tract is traditionally recognized as an organ for absorption. However, absorption in this organ does not occur without selection, which is based on several factors. The first factor is the size of the molecules. Large molecules are not absorbed unless their quantities are minute, such as in the case of antigen delivery (eg, polio vaccine). Hepatic, pancreatic, gastric, and intestinal enzymes will degrade typical macromolecules, such as proteins, polypeptides, and polysaccharides, into much smaller molecules. These smaller molecules are often hydrophilic and require the assistance of protein transporters in order for uptake into the cells to occur. This size restriction is enabled by the mechanical barrier of the intestinal cells, which is made of mucus, cell membranes, and tight junctions.

Several small molecules can diffuse into the circulatory system with ease via passive diffusion, but others cannot because of the presence of efflux transporters (eg, P-glycoproteins (P-gps)), which prevent the entrance of certain lipophilic substances. Small molecules that are taken up by the intestinal cells can be rapidly metabolized by the phase I and phase II enzymes in the enterocytes. Phase II enzymes appear to be the major metabolic enzymes that are capable of deactivating a large number of polyphenols and phenolics.[2] The excretion of phase I metabolites is usually straightforward, with the vast majority using the same pathway as the parent compound. The excretion of many phase II metabolites, especially those that are highly hydrophilic, requires the action of efflux transporters.[2] These transporters are often coupled to the phase II enzymes to achieve optimal functionality.

Finally, the expression of enzymes and transporters is highly region-dependent. Duodenal expression of an enzyme or a transporter is usually, if not always, different from those of the terminal ileum and colon.[3,4] These region-dependent expression levels can often be used as indicators that a transporter or enzyme is involved in drug uptake via experimental systems (eg, regional intestinal perfusion) that take advantage of this fact.

11.1.2 Hepatic barrier

The hepatic barrier is distinctly different from the intestinal barrier, in that the liver tends to be relatively uniform. The primary barrier of the liver is the presence of vast amounts of metabolic enzymes, which tend to inactivate lipophilic substances absorbed from the intestine that also escape intestinal metabolism. Both phase I and phase II enzymes are expressed in abundant quantities and varieties in the liver. The main phase I enzymes in the liver belong to the cytochrome P450 superfamily. These enzymes often produce metabolic products that can be further conjugated (via phase II enzymes) to hydrophilic moieties so that the final metabolites are highly hydrophilic. Then hydrophilic metabolites are often excreted to bile or carried to the kidneys for elimination.

Another important component of the liver barrier is the presence of various efflux transporters that line the bile canalicular membrane.[5] These transporters are capable of excreting parent compounds, phase I metabolites, and phase II metabolites out of the hepatocytes. Similar to the phenomena in the intestine, the functions of enzymes and efflux transporters can be interdependent, thereby resulting in significantly faster and more extensive metabolism than predicted. These coupling processes have been labeled as *double jeopardy* and *revolving door* mechanisms; explanations of these processes are provided later in the chapter.

11.2 PATHWAYS OF DRUG ABSORPTION

Absorption is the process of transferring chemical substances from the GI tract through its wall into the bloodstream and lymphatic stream. In general, absorption does not differ between drugs and food, but different mechanisms of absorption and differences in absorption rates occur in different parts of the GI tract. The absorption of drugs depends on physicochemical factors, including formulation factors (eg, rates of disintegration and rates of drug release from polymeric dosage forms), drug factors (eg, solubility and lipophilicity), and biological factors (eg, stomach-emptying rate and GI membrane permeability).

The two main mechanisms of drug transport across the GI epithelium are transcellular (ie, across cells) and paracellular (ie, between cells). The transcellular pathway is further divided into simple passive diffusion,

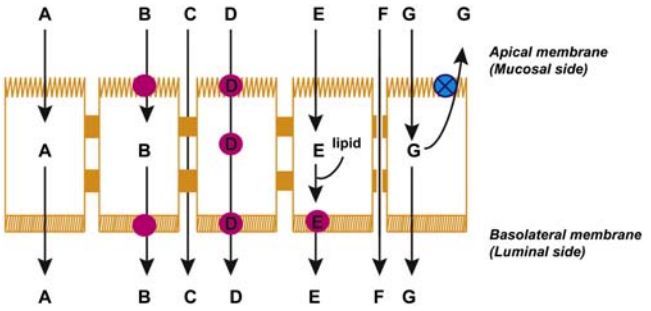


FIGURE 11.1 Potential modes of absorption of drug molecules across the intestinal epithelium: A, transcellular passive diffusion; B, carrier-mediated transcellular diffusion; C, paracellular diffusion; D, transcellular diffusion by endocytosis; E, transcellular diffusion and incorporation into lipid particles; F, paracellular diffusion with a modulator of tight junctions; G, transcellular diffusion modified by a polarized efflux mechanism.

carrier-mediated transport (active transport and facilitated diffusion), and endocytosis. These pathways are illustrated in Fig. 11.1. Transport across the healthy intestinal epithelium by the paracellular route (pathway C) is minimal because of the presence of tight junctions. Only small hydrophilic molecules are allowed to pass between cells unless a modulator of the tight junctions (pathway F) is present.[6,7] Transcellular transport of a molecule can occur via a passive mechanism (pathway A) or by a specific carrier (pathway B). For passive flux to occur, a molecule must have the correct physicochemical properties (eg, size, charge, lipophilicity, hydrogen-bonding potential, and solution conformation) to cross the lipophilic apical and basolateral membranes and diffuse through the cytoplasm, which is an aqueous environment separating the two membranes.[8] However, compounds that are absorbed by this route may be substrates for intracellular metabolism and apically polarized efflux mechanisms (pathway G), such as P-gps, which transport compounds from inside the cell back into the intestinal lumen.[8,9] Compounds may also be transcellularly absorbed by utilizing naturally occurring carriers that will transport the compounds from the lumen into the cell. Some drugs are absorbed via carrier-mediated pathways. For instance, L-dopa and oral cephalosporins are absorbed by amino acid and dipeptide transporters, respectively.[10,11] Endocytosis of compounds (pathway D) is minimal in the adult small intestine and is not a probable mechanism of drug absorption by the intestine. Most orally administered drugs are absorbed by passive transport.[8]

11.2.1 Paracellular diffusion

The paracellular pathway involves the transport of materials between cells rather than across cells. The pathway differs from all the other absorption pathways described later in this chapter. In the intestine, the cells are joined together via closely fitting tight junctions on their apical side. The intercellular spaces occupy only about 0.01% of the total surface area of the epithelium. The tightness of these junctions can considerably vary between different epithelia in the body. The paracellular pathway decreases in importance downstream of the GI tract because the numbers and sizes of the junctional gaps between the epithelial cells decrease.

The paracellular route of absorption is important for the transport of ions, such as calcium, and the transport of sugars, amino acids, and peptides at concentrations above the capacity of their transport carriers. Small hydrophilic and ionized drugs are not distributed into cell membranes, and they cross the GI epithelium via the paracellular pathway. The molecular weight cutoff for the paracellular route is usually considered to be 200 Da, although some larger drugs have been shown to be absorbed via this route, albeit usually at exceedingly small quantities. The paracellular pathway can be divided into a convective (solvent drag) component and a diffusive component. The convective component is the rate at which the compound is carried across the epithelium as the result of the water flux.

11.2.2 Passive diffusion

Passive diffusion is the preferred main route of absorption; thus, it is more likely to be used by low-molecular-weight lipophilic drugs. In the process, drug molecules pass the lipid membrane via passive diffusion from a region of higher concentration in the lumen to a region of lower concentration in the blood. This lower concentration is primarily maintained by the blood flow. The rate of transport is determined by the physicochemical properties of the drug, the nature of the membrane, and the concentration gradient of the drug across the membrane. The process initially involves the partitioning of the drug between the aqueous fluids within the GI tract and the lipoidal-like membrane of the epithelial membrane. The drug in the membrane diffuses across epithelial cells or cells within the GI barrier to the blood in the capillary network of the lamina propria. Upon reaching the bloodstream, the drug is rapidly carried away by blood flow, thereby maintaining a much lower concentration at the basal side than that at the apical side, thereby maintaining the so-called sink effect. If the cell membranes and fluid regions of the GI-blood barrier are modeled as a single membrane, the stages of GI absorption could be represented by the model in

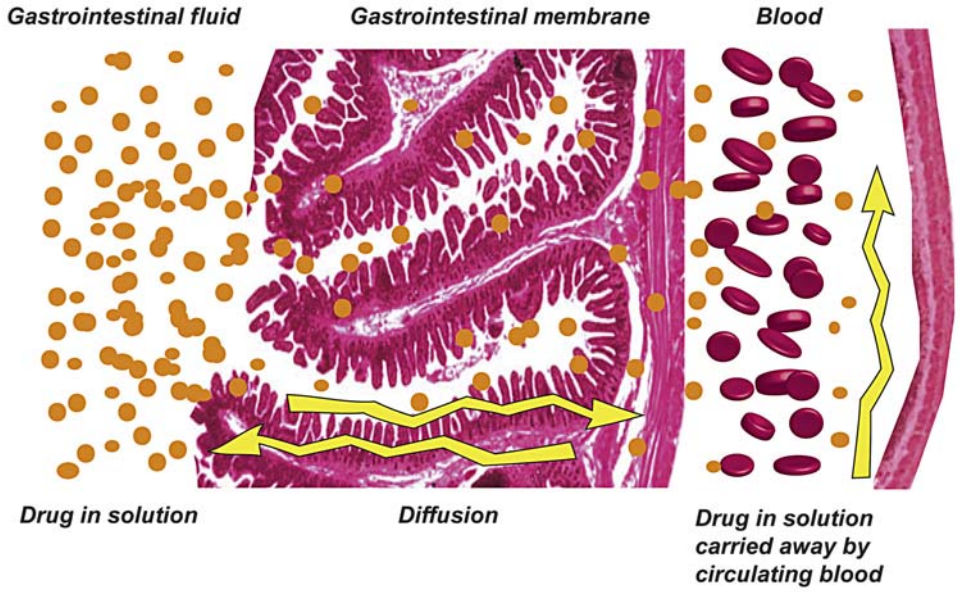


FIGURE 11.2 Diagrammatic representation of absorption via passive diffusion.

Fig. 11.2. The mechanism of the passive pathway follows Fick's law, where the absorption rate is proportional to the drug concentration gradient and the surface area. Given the notably different lipid and protein compositions of the apical and basolateral membranes, the rate of passage across these two barriers could be different for the same molecule.

The pathway has been assumed to exist solely in one single absorbable species. However, many drugs are weak electrolytes that exist in aqueous solution as two species: the unionized species and the ionized species. The unionized form of a weak electrolyte drug exhibits greater lipid solubility compared with the corresponding ionized form, and the GI membrane is more permeable to the unionized species. Thus, the rate of passive absorption of a weak electrolyte is related to the fraction of the total drug that exists in the unionized form in solution in the GI fluids at the site of absorption. This fraction is determined by the dissociation constant of the drug (eg, its pK_a) and by the pH of the aqueous environment. According to the pH-partition hypothesis, if the pH on one side of a cell membrane differs from the pH on the other side of the membrane, then the drug (weak acid or base) will be ionized to different degrees on the respective sides of the membrane. Moreover, the total drug concentration (ionized plus unionized drug) on either side of the membrane will be unequal, with the compartment in which the drug is more highly ionized containing the greater total drug concentration. For these reasons, a weak acid will be rapidly absorbed from the acidic environment (such as pH1.2), whereas a weak base will be poorly absorbed from the acidic environment.

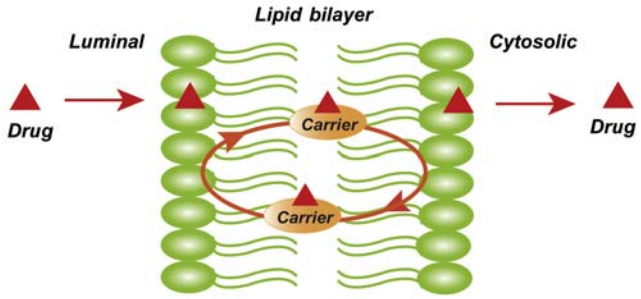


FIGURE 11.3 Schematic representation of the carrier-mediated transport of a drug across the intestinal membrane.

11.2.3 Carrier-mediated transport

Theoretically, a lipophilic drug may pass through a cell or move around cells during the absorption process. If the drug has a low molecular weight and is lipophilic, the lipid cell membrane is not an effective barrier to drug diffusion and absorption. For hydrophilic drug molecules, numerous specialized carrier-mediated transport systems, present in the body (especially in the intestine), may be used for their absorption. Normally, these carriers are utilized for the absorption of ions and nutrients and involves specific interaction between the molecule and the transporter or carrier in a saturable mechanism, as shown in Fig. 11.3. Two types of transport exist: active and passive carrier-mediated transport. The active carrier-mediated transport is either Na^+- or H^+-linked with bifunctional carriers. This process requires metabolic energy and can work against a concentration gradient of the substrate. By contrast, passive carrier-mediated

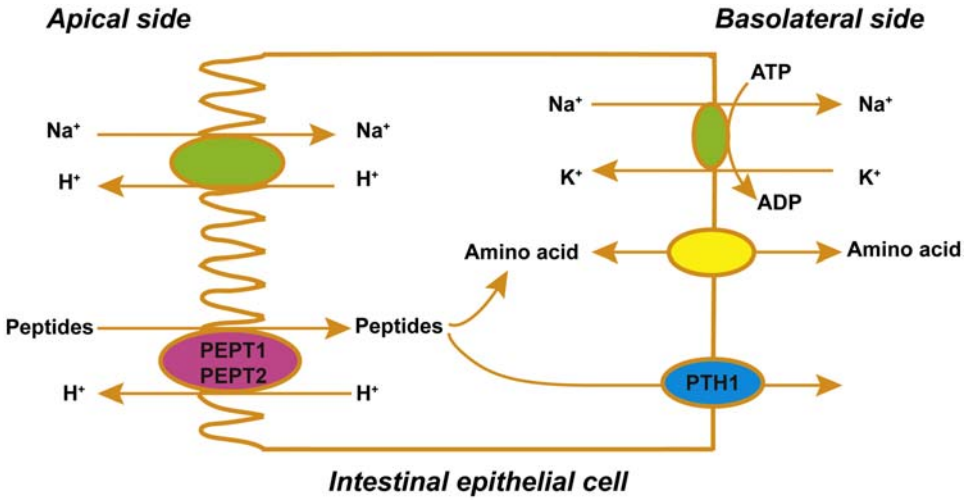


FIGURE 11.4 Accepted model for intestinal transepithelial peptide transport.

transport does not require metabolic energy and is driven by the concentration gradient of substrate. These carrier-mediated mechanisms are utilized by some drugs as pathways for entering cells. The best documented examples are β-lactams and cephalosporin antibiotics,[12] which are substrates for proton-driven dipeptide transporters. Peptide mimetic angiotensin-converting enzyme inhibitors, such as captopril, are also absorbed by carrier-mediated systems.[13]

11.2.3.1 *Active transport*

Compared with passive diffusion, active transport is an energy-consuming process where compounds can be transported against a concentration gradient across a cell membrane (ie, transport can occur from a region of lower concentration to one of higher concentration). This process is illustrated in Fig. 11.3. The energy arises either from the hydrolysis of adenosine triphosphate (ATP) or from the transmembrane ion gradient, electrochemical potential, or both. Numerous carrier-mediated active transport systems or membrane transporters are available in the small intestine and can be present on either the apical or basolateral membrane. These membrane transport systems include peptide transporters; nucleoside transporters; sugar transporters; bile acid transporters; amino acid transporters; organic anion transporters (OATs); efflux transporters of multidrug-resistant (MDR) proteins, such as MDR1, multidrug-resistant associated proteins (MRPs), and breast cancer resistant protein (BCRP); and vitamin transporters.

Unlike passive absorption, where the rate of absorption is directly proportional to the concentration of the absorbable species of the drug at the absorption site, active transport only proceeds at a rate that is proportional to the drug concentration at low concentrations. The carrier mechanisms may become saturated at higher concentrations, and further increases in drug concentration may not increase the rate of absorption (ie, the rate of absorption remains constant).

Competition occurs between two similar substrates for the same transport mechanism, and inhibition of absorption of one or both compounds is another characteristic of carrier-mediated transport. Some substances may be absorbed by simultaneous carrier-mediated and passive transport processes. Generally, the contribution of the carrier-mediated process to the overall absorption rate decreases with concentration and is negligible at a sufficiently high concentration.

In summary, the characteristics of active transport mechanisms are as follows: (1) The process must have one or multiple carrier transporters, (2) the process must be activated by ATP energy, (3) the process shows temperature dependence, and (4) active transport can be competitively inhibited by substrate analogs or specific inhibitors. Active transport also works to excrete several drugs in the intestine, kidney, liver, and bile.

11.2.3.1.1 Peptide transporters

As described previously, most oral chemical drugs rely on passive diffusion to cross cell membranes. However, some orally administered peptidomimetic drugs are absorbed through the intestinal peptide transport system[14,15] (Fig. 11.4). In this transport process, the proton gradient provide the driving force for peptide uptake into intestinal epithelial cells via proton-dependent peptide transporters located in the apical membranes.[16] Peptides that are resistant to hydrolysis by intracellular peptidases are transported across the basolateral membrane via the basolateral peptide transporters.[17] The Na^+/H^+-exchanger generates and

maintains the inward proton gradient on the luminal surface, while the Na^+/K^+-ATPase present in the basolateral membrane maintains a low intracellular sodium concentration. Protons are cotransported with peptides across the epithelial membrane; thus, this system is also referred to as the H^+-dependent peptide cotransport system.[16] Peptide transporters can accept dipeptides and tripeptides as physiological substrates, thereby indicating their much broader substrate specificity than other nutritional transporters. Consequently, drugs and xenobiotics with structures resembling small peptides (eg, β-lactam antibiotics), are recognized by the peptide transporters. Therefore, peptide transporters work as nutritional transporters and drug transporters (Fig. 11.4).

Currently, the peptide transporters are divided into two types: those localized at the brush-border membranes of epithelial cells, and those localized at the basolateral membrane of intestinal and renal epithelial cells. Two-border-type peptide transporters have been identified and designated as PEPT1 and PEPT2. These polytopic integral membrane proteins have 12 predicted membrane-spanning domains (MSDs) and NH_2- and COOH-terminal ends facing the cytosol. Western blot analysis of protein preparations of the intestine and kidney have identified PEPT1 and PEPT2 via immunoreactivity as glycosylated proteins with molecular masses of about 75 and 100 kDa, respectively.[18,19] The mammalian PEPT1 protein is comprised of 701−710 amino acids, depending on the species, and is highly glycosylated.[20−22] Digestion of PEPT1 (about 75 kDa) with endoglycosidase A shifts its mass significantly to 63 kDa.[22] The PEPT2 gene encodes a 729−amino acid protein[19] and shares 48% amino acid identity with PEPT1 predominantly in transmembrane domains (TMDs). PEPT1 and PEPT2 share the lowest identity in the large extracellular loop connecting TMD9 and TMD10, perhaps signaling some difference in substrate-binding recognition. The PEPT2 protein is also glycosylated; the molecular mass of the mature protein is about 107 kDa and its nonglycosylated mass is 83 kDa.[23] The genomic organization of peptide transporter genes has also been elucidated. Human PEPT1 and PEPT2 genes, for example, have been mapped to chromosomes 13q33-34 and 3q13.3-q21, respectively.[21,24]

Peptide/histidine transporter 1 (PHT1) have substrate specificity similar to that of PEPT1 and PEPT2 but possess other distinct functional properties. Unlike PEPT1 and PEPT2, these transporters are localized at basolateral membranes of intestinal and renal epithelial cells.[25] Absorption of peptidelike drugs through the intestinal epithelial requires the crossing of two distinct membranes; that is, uptake by epithelial cells from the lumen across the brush-border membrane (presumably by PEPT1/PEPT2), followed by transfer to the bloodstream across the basolateral membrane (presumably by PHT1). Although orally active β-lactam antibiotics are efficiently taken up from the intestine, these compounds are difficult to move across the basolateral membranes by passive diffusion because of their hydrophilicities. This leads to the study of the basolateral peptide transporter in the intestinal epithelial cells. The characterization of peptidelike drug transport via the basolateral membranes of Caco-2 cells grown on microporous membrane filters demonstrated that the PHT1 transporter, which is distinguished from PEPT1, is expressed in the basolateral membranes of intestinal epithelial cells.[26,27]

The intestinal peptide transporters, especially PEPT1, have been key target molecules for prodrug approaches.[28] Under this scheme, appropriately designed molecules in the form of dipeptide or tripeptide analogs can be absorbed across the intestinal brush-border membranes via PEPT1 and may be absorbed intact or hydrolyzed intracellularly by peptidases or esterases prior to exiting the cells. In addition to PEPT1, the intestinal basolateral PTH1 transporter has the ability to transport nonpeptidic compounds such as δ-ALA, which is a precursor of porphyrins and heme. Here, δ-ALA plays an important role in the production of heme-containing proteins.[27] The transport of δ-ALA by PEPT1 and the basolateral peptide transporter has been reported to explain the good bioavailability of δ-ALA (approximately 60% in humans).[27,29]

11.2.3.1.2 Amino acid transporters

The endothelial lining of blood vessels provides a barrier for the exchange of nutrients and is actively involved in the local control of vascular homeostasis. Multiple transport systems mediate the influx of cationic, neutral, sulfonic, and anionic amino acids across the plasma membrane of mammalian cells. Molecular cloning approaches have led to the identification of Na^+-dependent and Na^+-independent amino acid transport systems. Table 11.1 summarizes the nomenclature used to designate these amino acid transport systems. Note that the ionic dependency and K_m corresponding to the different transport systems were compiled from the cited publications and do not necessarily reflect the transport properties of all cell types.[30]

The different carrier proteins mediating transport of cationic amino acids include the Na^+-independent systems y^+, y^+L, $b^{0,+}$, and b^+, as well as the Na^+-dependent system $B^{0,+}$. System b^+ was originally described in mouse blastocysts and is highly specific for cationic amino acids, whereas the other systems also can transport neutral amino acids. System y^+ is the principal cationic amino acid transport system expressed in NO-producing cells and most likely plays a key role in regulating L-arginine supply for NOS. Although

TABLE 11.1 Mammalian Amino Acid Transport Systems and Identified Transporters[30]

Transporter system	Associated protein or cDNA	SLC number	Amino acid selectivity	Ionic dependency	K_m, μM
CAT-ASSOCIATED TRANSPORT					
y^+	CAT-1	SLC7A1	Cationic	No	70–250
	CAT-2A or CAT-2/2α	SLC7A2	Cationic	No	2150–5200
y^+	CAT-2B or CAT-2/2β	SLC7A2	Cationic	No	38–380
$y^{+?}$	CAT-3	SLC7A3	Cationic	No	40–120
	CAT-4	SLC7A4	Cationic (Arg)	No	450–910
HETERODIMERIC GLYCOPROTEIN (4F2HC, CD98)–ASSOCIATED TRANSPORT					
y^+L	y^+LAT1	SLC7A7	Neutral/cationic	Na^+ (neutral)	340
y^+L	y^+LAT2	SLC7A6	Neutral/cationic	Na^+ (neutral)	6–10
L	LAT1/LAT2	SLC7A8	Large neutral	No	30–300
x_c^-	xCT	SLC7A11	Cystine/glutamate	No	40–92
Asc	Asc-1	SLC7A10	Small neutral	No	9–23
rBAT-ASSOCIATED TRANSPORT					
$b^{0,+}$	$b^{0,+}$, AT	SLC7A9	Neutral/cationic cystine	No	88
UNKNOWN HEAVY CHAIN–ASSOCIATED TRANSPORT					
asc	Asc-2	SLC7A10	Small neutral	No	2.9
NA^+ AND/OR CL^- OR H^+–ASSOCIATED TRANSPORT					
A	ATA1/2/3(SAT1/2/3 or SA 1/2/3)	SLC38A1	Neutral/N-Me group	Na^+	> 200
N	SN 1/2/3	ASNS	Gln, Asn, His	Na^+/H^+	150–1600
ASC	ASCT 1/2[a]	SLC1A4	Neutral	Na^+	9–464
$B^{0,+}$	$ATB^{0,+}$	SLC6A14	Neutral/cationic	Na^+/Cl^-	23–140
Pro	PROT[a]	SLC6A7	Praline	Na^+/Cl^-	6
Gly	GLYT 1/2[a]	SLC6A5/9	Glycine	Na^+/Cl^-	17
Betalike	mTAUT[a]	SLC6A6	β-amino-acids	Na^+/Cl^-	3–13
	GAT-1-3	SLC6A11/12	GABA and/or β-amino acid	Na^+/Cl^-	1–20
	BGT-1				
	EAAT1-5[a]	SLC1A3/2/1/4/5	Glu and/or Asp	Na^+, OH^-/HCO_3	18–97
X^-_{AG}			Glu, Asp	Na^+	About 200
x^-_{AG}			Glu	No	About 230

[a]*Cloned high-affinity amino acid transporters from the brain or retina that may not exhibit kinetic properties described in other peripheral cell types. Lowercase nomenclature for the classical transport systems denotes Na^+-independence.*

Summary of the classical nomenclature for the majority of amino acid transport systems and the nomenclature adopted for recently cloned amino acid transport proteins. The substrate specificity, ionic dependency, and Michaelis constant (K_m) values included in this table were compiled from the cited references and reviews.

Ala, alanine; Ser, serine; Cys, cysteine; Gln, glutamine; Asn, asparagine; His, histidine; Glu, glutamate; Asp, aspartate.

information on ASCT-, PROT-, GLYT-, TAUT-, and EAAT-associated amino acid transport in vascular cells is limited, recent evidence indicates that bovine aortic endothelial cells express a taurine transporter sharing a high degree of sequence homology with that of the mTAUT complementary DNA (cDNA) isolated from the brain.[31]

Similarities between the transport of cationic amino acids in renal proximal tubules and the small intestine have been recognized. In patients with the renal abnormality cystinuria, lysine and ornithine transport in the intestinal tract is also defective. For this reason, and with considerable prescience, the genetic defect underlying cystinuria is also evident in the small intestine. This

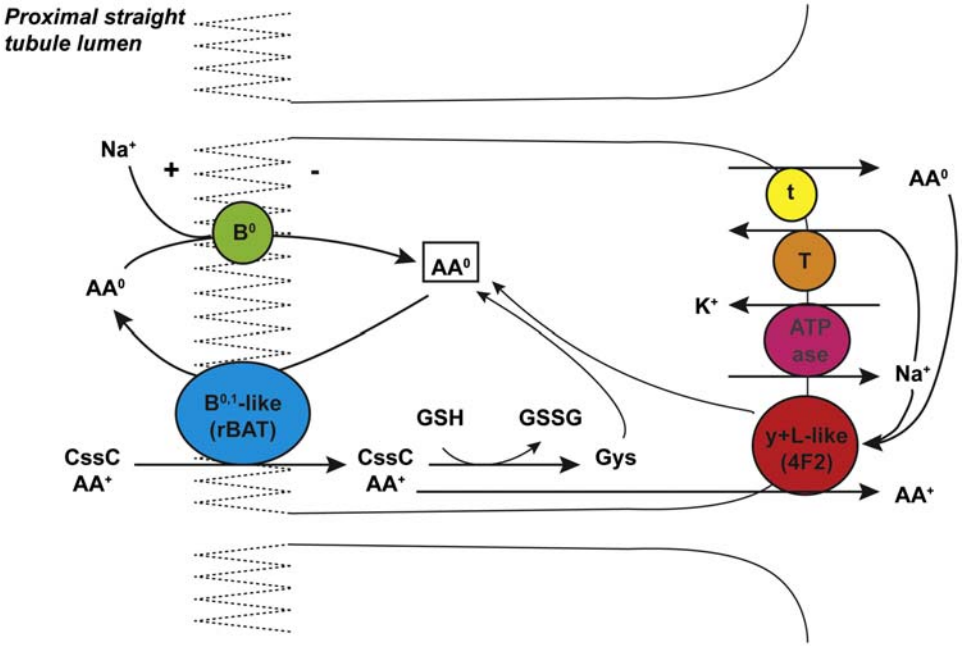


FIGURE 11.5 Model for renal reabsorption of cationic amino acids (AA^+) and cysteine (CssC). Participation of various broad-scope systems ($b^{0,+}$, y^+L, and $B^{0,+}$) and interactions with neutral amino acids (AA^0) are indicated. Note the proposed functions of rBAT ($b^{0,+}$) and 4F2 (y^+L) proteins in apical and basal membranes, respectively. *GSH*, glutathione; *GSSG*, oxidized glutathione. *Source: Redrawn from Deves R, Boyd CA. Transporters for cationic amino acids in animal cells: discovery, structure, and function.* Physiol Rev *1998;78:487–545.*

notion is compatible with the findings of early studies, which indicated that net transport of three basic amino acids against a concentration gradient occurs in the hamster intestine in vitro, and that these amino acids share the same transport system as cysteine.[32]

The Na^+-dependent and Na^+-independent transport systems for cationic amino acids have been identified in the brush border of the small intestine. Evidence also shows that efflux across the basolateral membrane constitutes a rate-limiting step of transepithelial cationic amino acid transport. During lysine transport across the intact epithelium, the intracellular amino acid concentration may be up to fivefold higher than that in the lumen; by contrast, the accumulation of solute in tissues is much less for neutral amino acids.[33,34]

The principles of transepithelial cationic amino acid transport exemplified in the intestine are also applicable to the kidney. In the kidney, the Na^+ dependence and electrogenicity of the brush-border transport of cationic amino acids are controversial. Thus, various studies with membrane vesicles conclude that Na^+-dependent secondary active transport occurs,[35] that transport is electrogenic but Na^+-independent,[36] and that electroneutral transport takes place in the absence of Na^+.[37] A model for the reabsorption of cationic amino acids in the kidney was recently proposed and is shown in Fig. 11.5.[32]

11.2.3.1.3 Organic anion-transporting peptides

OATs and organic cation transporters (OCTs) are two of the major transporter families responsible for the absorption, distribution, and elimination of environmental xenobiotics, plant or animal toxins, and drugs. Although both transport systems have distinct substrate preferences, OAT is for those carrying negative charges and OCT is for those carrying positive charges. Both share several common features, including similar predicted transmembrane topology, interaction with numerous structurally and pharmacologically diverse compounds, and broad expression patterns in tissues, such as kidney, liver, intestine, brain, and placenta. Six members of the OAT family have been identified thus far: OAT1, OAT2, OAT3, OAT4, OAT5, and URAT1. These OATs have primary structures of 526–568 amino acid residues and a predicted secondary structure of 12 TMDs. These multispecific transporters are expressed in a variety of tissues, including the kidney, liver, brain, and placenta, with a wide range of recognized substrates. Six members of the OCT family have been identified: OCT1, OCT2, OCT3, OCTN1, OCTN2, and OCTN3. Based on their primary structures, OCTs and OCTN3 are distinct subfamilies within the OCT group. Like OATs, common structural features of OCT family members include 12 TMDs, a large hydrophilic loop between TMDs 1 and 2 carrying several potential glycosylation

sites, and multiple potential phosphorylation sites localized on the large intracellular loop between TMDs 6 and 7.[38]

OAT1 and OAT3 transporters belong to the classically described OATs in the kidney proximal tubules, which utilizes a tertiary transport process to move organic anions across the basolateral membrane into the proximal tubule cells for subsequent exit/elimination across the apical membrane into urine. This tertiary process in the basolateral membrane includes the following steps:

1. Na^+ K^+-ATPase pumps Na^+ out of the cells and generates an inwardly direct Na^+ gradient.
2. This Na^+ gradient then moves dicarboxylates into the cells through an Na/dicarboxylate cotransporter ($SDCT_2$). This process, in conjunction with an ongoing cellular metabolic activity, sustains an outwardly directed dicarboxylate gradient.
3. The dicarboxylate gradient transfers organic anions into the cells via OAT1 and OAT3, which serve as organic anion/dicarboxylate exchangers.

Whether other members of the OAT family, such as OAT2 and OAT4, use a similar transport mechanism remains unknown. Oat5 is exclusively localized in the kidney. Oat5 urinary excretion was recently proposed as a potential early biomarker of acute kidney injury.[39]

After being transported across the basolateral membrane into the proximal tubule cells, organic anions exit across the apical membrane into the urine for elimination. However, the mechanism for this apical step is not yet well understood.[40,41]

The characteristics of organic cation transport by OCT1, OCT2, and OCT3 are electrogenicity, reversibility of transport direction, and independence from Na^+.[42] For OCTNs, the transport modes depend on the substrates; and hOCTN1 and rOCTN1 mediate pH-dependent and Na^+-dependent transport of trimethylamine or TEA.[43,44] However, mOCTN1 mediates carnitine transport in an Na^+-dependent manner.[45] OCTN2, which has significant affinity for carnitine and TEA, transports carnitine in an Na^+-coupled manner. However, OCTN2 transports other organic cations, including TEA, in an Na^+-independent manner.[46,47] OCTN2-mediated transport is also pH-dependent and electrogenic, as was recently demonstrated by the two-electrode voltage clamp technique.[48] By contrast, the OCTN3-mediated transport of carnitine is Na^+-independent.[45]

Three synthetic opioid peptides: DPDPE, DADLE, and deltorphin II are transported by certain members of the OATP/Oatp family, including rodent Otap1, Oatp2, and Oatp4[49,50] and human OATP-A, OATP-C, and OATP-8.[50–53] Enkephalins may also be recognized by some of these transporters. The primary location of most of these transporters is the basolateral membrane of hepatocytes.[49,51,54–56] Therefore, these transporters play a critical role in the uptake of opioid peptides by the liver for subsequent elimination into bile. OATP-A is located in the brain capillary endothelial cells in humans[50,53] and may be expected to be an important determinant of the transfer of opioid peptides across the blood–brain barrier. Evidence of marked changes in the ability of OATP-A to transport opioid peptides due to genetic polymorphisms has been reported.[53]

11.2.3.2 Facilitated transport

Facilitated transport is a carrier-mediated process that differs from active transport because the facilitated process cannot transport a substance against its concentration gradient. Therefore, facilitated transport does not require energy input but does require a concentration gradient as the driving force, which is similar to passive diffusion. In facilitated diffusion, substances are transported down the concentration gradient at a much faster rate than would be anticipated based on the molecular size and polarity of the molecule for pure diffusion. The facilitated process, like active transport, is saturable and subject to inhibition by competitive inhibitors. Facilitated diffusion seems to play a very minor role in drug absorption.

11.2.3.2.1 Nucleoside transporters

Nucleosides are precursors for nucleic acid synthesis, which is fundamental to the control of growth and metabolism in all living systems. Nucleosides are hydrophilic and diffuse slowly across cell membranes. To facilitate the salvage of these precursors for nucleotide biosynthesis, cells employ a complex transport systems consisting of multiple carrier proteins known as the *nucleoside transporters*. The physiologic roles of these transporters include the salvaging of precursors for DNA and RNA synthesis and the regulation of adenosine-mediated neuromodulation by controlling its concentrations at the receptor site. Several therapeutic nucleoside analogs rely on nucleoside transporters to enter or exit cells. Consequently, the expression and functional characteristics of these transporters in the absorptive/excretory organs and target/nontarget cells exert an important effect on the pharmacokinetics, efficacy, and toxicity of nucleoside analogs.[57]

Two major classes of nucleoside transport systems, equilibrative and concentrative, have been identified. The equilibrative system comprises facilitated carrier proteins, whereas the concentrative system consists of Na^+-dependent secondary active transporters. Multiple subtypes with distinct substrate specificity and inhibitor sensitivity exist within each major class.

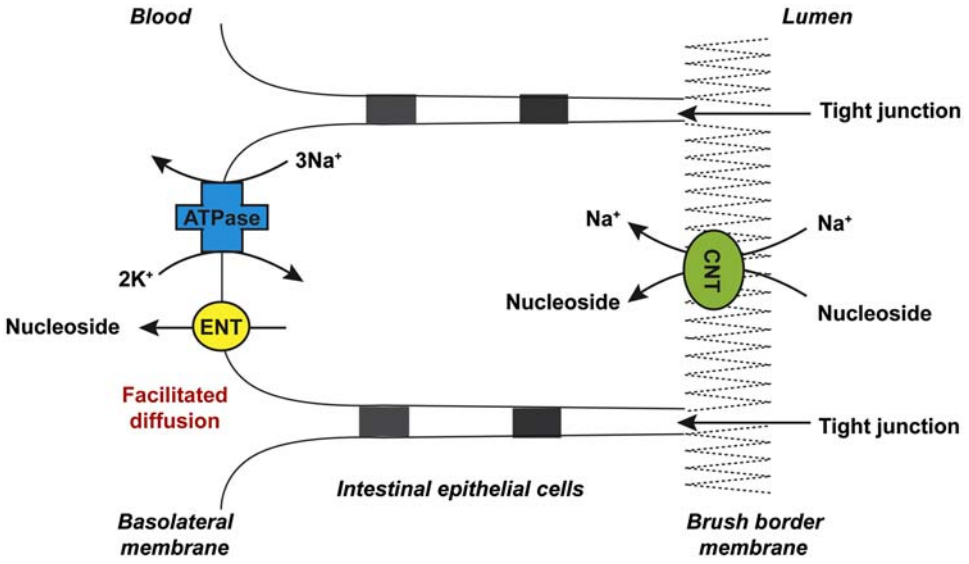


FIGURE 11.6 Model of transepithelial nucleoside flux mediated by Na^+-dependent nucleoside transporters (CNTs) and ENTs in the lumen of the intestine.

Cloned equilibrative nucleoside transporters (ENTs) share significant sequence homology and belong to a so-called ENT gene family, whereas genes encoding Na^+-dependent nucleoside transporters belong to a concentrative nucleoside transporter (CNT) gene family.

In polarized cells, such as renal and intestinal epithelia and hepatocytes, ENTs and CNTs are asymmetrically distributed between the apical and basolateral membranes to mediate the vectorial-transepithelial flux of nucleosides (Fig. 11.6). ENTs mediate the facilitated diffusion of nucleosides across membranes and bidirectionally function according to the substrate concentration gradient. Based on sensitivity to nitrobenzylthioinosine (NBMPR), which is a tight-binding and highly specific inhibitor, ENTs have been classified into two subtypes: *es* (equilibrative sensitive) and *ei* (equilibrative insensitive). The *es* subtype binds to NBMPR with high affinity via noncovalent interactions at a high-affinity binding site on the transporter protein. By contrast, the *ei* subtype is not affected by NBMPR at nanomolar concentrations and becomes inhibited only at high NBMPR concentrations.[58] Both *es* and *ei* transporters exhibit broad substrate selectivity and are widely distributed in different cell types and tissues. More recently, the two transporters were renamed as ENT1 and ENT2 according to their characteristics and belong to a large NET gene family, the members of which are also found in yeast, protozoans, insects, and plants.[59]

Moreover, two new mammalian ENT members (ENT3 and ENT4) with unknown function have been identified from the public sequence database. Compared with ENTs, which mediate the passive downhill flux of nucleosides, CNTs mediate the active uphill transport of nucleosides. These transporters utilize the inwardly directed Na^+ gradient established by the ubiquitous Na^+/K^+-ATPase to move substrates into cells against their concentration gradients. Unlike ENTs, Na^+-dependent nucleoside transporters exhibit distinct substrate selectivity for purine or pyrimidine nucleosides to regulate various physiological processes, including neurotransmission and cardiovascular activity.

Based on functional studies in cell and tissue preparations, five Na^+-dependent nucleoside transporter subtypes (N1–N5) have been identified. The N1 (or *cif*) subtype is mainly purine-selective but also transports uridine. The N2 (or *cit*) subtype is pyrimidine-selective but also transports adenosine. The N3 (or *cib*) subtype is broadly selective and transports both purine and pyrimidine nucleosides. The N4 subtype transports pyrimidine and guanosine and is characterized in brush border membrane vesicles isolated from the human kidney. The N5 subtype, which is NBMPR sensitive and transports formycin B and cladribine, is described only in human leukemia cell lines. To date, the genes encoding N1–N3 transporters have been cloned, and these genes belong to a CNT gene family. The genes encoding N4 and N5 transporters have not been identified.[57]

The clinical significance of nucleoside transporters can be summarized as follows:

- A number of anticancer nucleoside analogs rely on nucleoside transporters to enter cells to reach their cellular targets. For these drugs, the expression of nucleoside transporters on cancer cells is a

prerequisite for the cytotoxicity of nucleoside analogs. Loss of nucleoside transporters in tumor cells can lead to drug resistance.
- If a drug is a substrate for nucleoside transporters, the distribution of these transporters in various tissues and organs (particularly absorptive and excretory organs) may influence its pharmacokinetic and toxicological properties.
- Nucleoside transporters themselves could serve as drug targets.

In conclusion, nucleoside transporters are critical in determining the intracellular bioavailability and systemic disposition of therapeutic nucleoside analogs. Knowledge of nucleoside transporters is important in the evaluation and prediction of the kinetics, targeting, efficacy, and toxicities of nucleoside-derived drugs.

11.3 PATHWAYS OF DRUG METABOLISM

Drug metabolism usually involves a variety of enzymes, which by definition (from a biochemical perspective) typically helps increase efficiency (by acting as a catalyst) to biotransform a substrate into a product by reducing the required threshold energy for the reaction. Two categories of drug metabolism are well recognized in the field today—namely, phase I and phase II metabolism. In general, phase I metabolism acts as the first step in the drug metabolic process, in which its product often becomes a substrate for phase II metabolism. When lipophilic endobiotics and xenobiotics enter cells through passive diffusion or uptake transporters, phase I enzymes often facilitate the first enzymatic reaction that acts to clear them from the body. The phase II reactions are the true detoxification pathways and produce metabolites that are generally highly water-soluble and can be readily excreted from the body in the form of bile or urine.

11.3.1 Phase I metabolism

Phase I drug-metabolizing enzymes (DMEs) can catalyze a wide array of chemical reactions that facilitate xenobiotic metabolism and elimination. A common theme apparent for all phase I enzymatic reactions is that the final products are usually modified to contain a chemically reactive functional group (ie, hydroxyl, amine, sulfa, and carboxylic acid).[60]

This property supports the well-documented concept that phase II metabolism (ie, glucuronidation, glycosidation, sulfation, methylation, acetylation, glutathione conjugation, amino acid conjugation, fatty acid conjugation, and condensation) is the true detoxification step for xenobiotic elimination in the body (discussed in detail next). Because phase I reactions add functional groups onto a substrate for phase II metabolism to occur, phase I metabolism becomes a necessary prerequisite for the detoxification of xenobiotics.

Phase I metabolism also occurs during nutrient digestion, which typically involves hydrolysis reactions. These reactions include the hydrolysis of proteins, polysaccharides, and polymeric nucleotides and nucleosides[61,62] and typically provide nutrients useful to the human body. Additional phase I reactions include hydrolysis of the ester bond, which is necessary for the processing of fatty acids and their derivatives, to produce monoglyceride and fatty acids for absorption.[63] We emphasize here that some drugs may also undergo hydrolysis in the GI tract and the liver, but these drugs are uncommon.

11.3.1.1 Oxidative metabolism

Many drugs and other xenobiotics are biotransformed by oxidative reactions, which are defined as reactions that remove electrons from a molecule. Two types of enzymatically mediated oxidation reactions are known:

- Oxygen is incorporated into the drug molecules by the insertion of a hydroxyl group (−OH). Hydroxylation is the most common reaction type in phase I metabolism and usually produces a chemically stable and more polar hydroxylated metabolite than the drug.
- Hydroxylation sometimes creates unstable compounds that spontaneously split into small molecules by losing parts of the drug molecules. Different biotransformation processes, such as dealkylation and deamination, are mediated by this basic mechanism.

When most experts in the drug metabolism field are asked about phase I metabolism, typical conversations focus on oxidation metabolism involving the cytochrome P450 (CYP) enzyme family because CYP oxidation reactions are very important and highly diverse; thus, these reactions have received a considerable amount of research focus. However, oxidation reactions not related to CYP are also observed and typically involve the metabolism of endogenous substrates. The list of common names for oxidative metabolism via hydroxylation include dehydrogenases (eg, alcohol dehydrogenase), oxidases (eg, amine oxidase), and aromatases (eg, P450arom).[64]

11.3.1.1.1 Cytochrome P450 enzymes

The majority of research on phase I metabolism has focused on cytochrome P450 (CYP) proteins (the number 450 reflects the absorption maxima of the enzymes

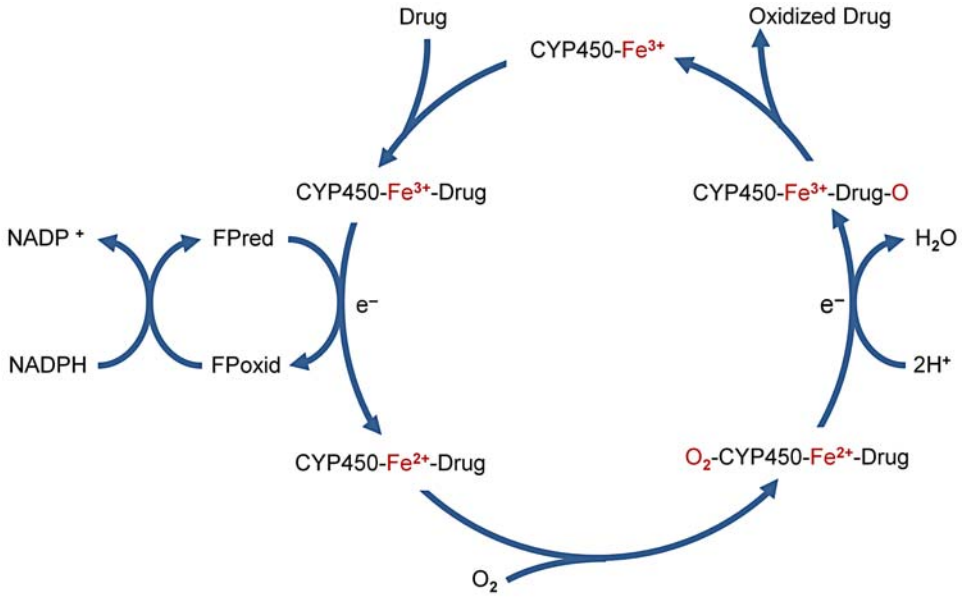


FIGURE 11.7 Generalized catalytic cycle for cytochrome P450-mediated oxidation.

at 450 nm using the CO-binding spectrum). As a superfamily of enzyme proteins, CYP plays a very important role in the phase I oxidative metabolism of drugs.

Because CYPs are responsible for the primary metabolism of a large variety of drugs in vivo,[65] determination of the mechanisms of CYP reactions has been emphasized. The importance of CYP can be observed from various perspectives. From a basic research perspective, CYP eliminates xenobiotics by preparing a substrate for phase II metabolism through addition of reactive functional groups. However, this does not mean that phase I metabolites cannot be excreted or eliminated from the body through either active or passive efflux. From the point of view of drug design, a thorough understanding of CYP behavior and its activities with xenobiotics can help predict possible adverse reactions of drugs and provide solutions to improve drug efficacy and administration.[66] Therefore, the US Food and Drug Administration (FDA) has paid a significant amount of attention to this subject. In a recent draft from the agency providing guidance on drug interaction studies, the need to define the metabolism of a new investigational drug clearly during early drug development was stressed.

CYPs are located in the endoplasmic reticulum, which can oxidize various xenobiotics. Because CYPs have a prosthetic group known as *iron protoporphyrin IX*, they are categorized as heme-containing enzymes. Heme iron is bonded to the rest of the enzyme by four pyrrole nitrogen atoms of protoporphyrin IX and two axial ligands to the porphyrin plane. The two axial ligands (known as the *fifth* and *sixth ligands*) and their complex configuration states enable CYPs to hydroxylate a wide variety of substrates. The overall main catalytic function of CYPs is to insert one atom of oxygen into a substrate (oxygenation) (Fig. 11.7).

11.3.1.1.2 Nomenclature of CYP

The nomenclature for CYP is essentially based on the degree of structural similarity among different enzyme isoforms. Current nomenclature guidelines suggest that members of CYP families share $\geq 40\%$ amino acid identity, while CYP enzymes in the same subfamily share $\geq 55\%$ amino acid identity. Individual enzymes within the same subfamily are identified based on their catalytic specificity, which is often defined using a so-called probe substrate. Virtually all CYP enzymes are designated with the root *CYP*, followed by an Arabic numeral for the gene family (eg, CYP3), a capital letter for the subfamily (eg, CYP3A), and another Arabic number for a particular gene (eg, CYP3A4).

Because of the wide variety of CYPs that exist in nature, the classification of these enzymes is also a rather complex issue because newly found isoforms are continuously published in the literature. Amidst this confusion, Dr David R. Nelson developed and maintained an updated website that reports all known CYP genes across species. In addition to this database, sequence alignments and information about known structures are also available on the Internet (http://drnelson.uthsc.edu/CytochromeP450.html). Humans have 57 genes and over 59 pseudogenes divided into 18 families and 41 subfamilies of CYP genes (Fig. 11.8). Most drugs are metabolized by CYP families 1, 2, and 3.

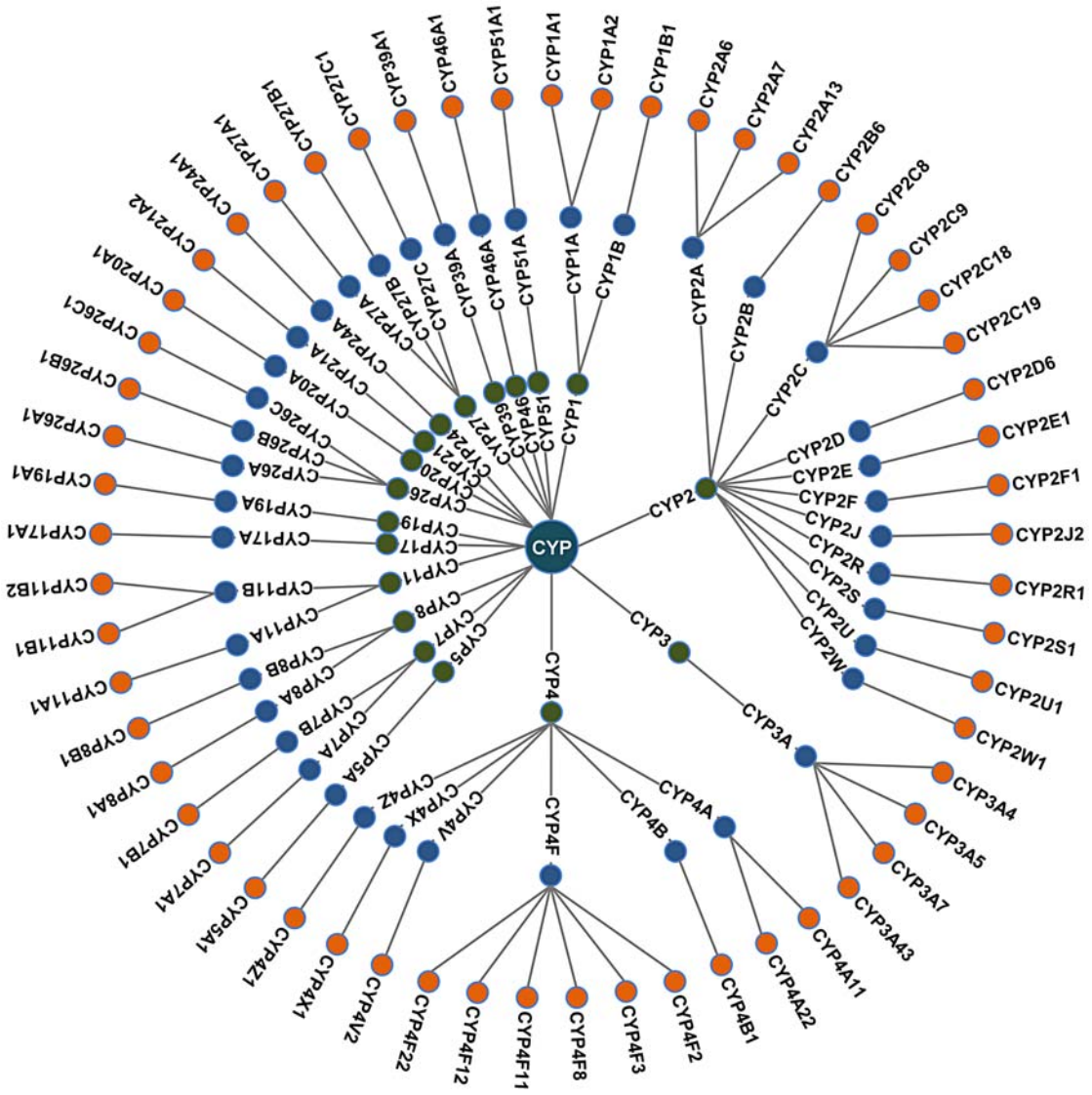


FIGURE 11.8 Phylogenetic tree representing human CYP proteins.

Despite the highly diverse amino acid compositions and enzyme mutations (commonly point mutations, such as single nucleotide polymorphisms) of CYPs that occur quite frequently in nature, the structures of CYPs are so well conserved that their catalytic functions can be well preserved. However, some mutations can cause serious damage to the catalytic capacity of CYP enzymes, leading to complicated pathologies such as Antley-Bixler syndrome.[60] Another example in which polymorphism greatly affects patient reaction to drugs is seen in the case of warfarin (ie, the S enantiomer form has higher activity than the R form). CYP2C9 is the main isoform responsible for S-warfarin metabolism. Patients with polymorphisms, particularly CYP2C9*2 and CYP2C9*3 (both heterozygous and homozygous), have major deficiencies in S-warfarin metabolism, making them susceptible to a higher incidence of life-threatening bleeding.[67]

11.3.1.1.3 Hydroxylation

Hydroxylation reactions are typically very common reactions for compounds that have aromatic rings in their chemical structures. Because many drugs and xenobiotics consist of aromatic rings, hydroxyl derivatives are quite commonly seen in nature for both drugs and xenobiotics. Another very common hydroxylation reaction is seen in aliphatics. Of interest to cancer research are epoxidation reactions where unstable intermediates are formed. Although sometimes derived from polycyclic compounds, epoxide formation is important, as it is often the first step in

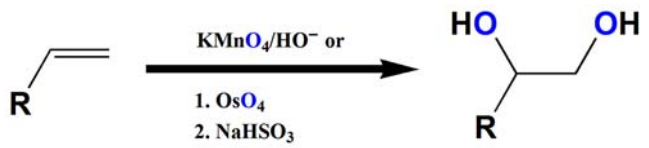


FIGURE 11.9 Chemical equation of alkene oxidization from donated hydroxyl groups.

aromatic hydroxylation, which may lead to carcinogen activation.

The essential formula that underlies all hydroxylation reactions is seen as chemically modified with the addition of one or more hydroxyl groups (−OH). In the example given in Fig. 11.9, alkene (starting compound) is oxidized through the addition of hydroxyl groups.

11.3.1.2 *Reductive metabolism*

Reductive metabolism in mammalian species is quite rare and seldom occurs in the metabolism of drugs and xenobiotics. However, reductive metabolism and ring fission frequently occur in the intestinal microflora.[68] The essential formula that underlies all reduction reactions is seen when a chemical is modified in two ways: (1) addition of hydrogen and (2) removal or reduction of oxygen. The commonality that exists for reductive metabolism is that it requires nicotinamide adenine dinucleotide phosphate (NADPH) as a necessary component in the reaction mixture but is generally inhibited by oxygen. This observation differs from mixed-function oxidase reactions, where oxygen is required for reaction.

The major groups of compounds metabolized by reductive reaction are compounds with nitro or azo functional groups. These compounds are reduced to their amines—for example, conversion of 3′-azido-3′-deoxythymidine (zidovudine) to 3′-amino-3′-deoxythymidine (Fig. 11.10). Aldehydes and ketones are also reduced to their corresponding alcohols by aldoketoreductases (Fig. 11.11). Reductive dehalogenation of halothane to 2-chloro-1,1-difluoroethylene (Fig. 11.12) is another important reductive reaction in the mammalian liver, especially under anaerobic conditions.

11.3.1.3 *Hydrolysis*

11.3.1.3.1 Necessity of hydrolysis

The necessity for hydrolysis with regard to both absorption and metabolite conversion (to parent compound) can be exemplified by the intestinal disposition pathways for compounds such as isoflavones. Natural isoflavones exist predominantly in the form of glycosides. In order for them to be absorbed, the lactase phlorizin hydrolase that resides on the surface of the intestinal lumen must be utilized to cleave the sugar

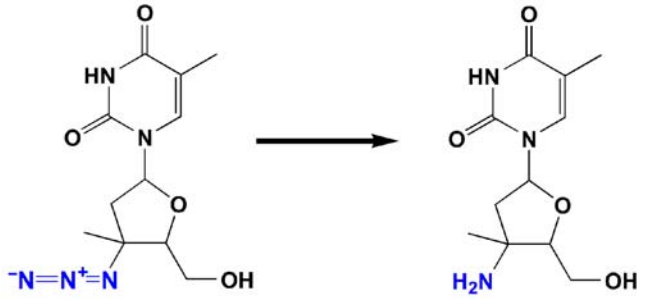


FIGURE 11.10 Chemical equation of azo-reduction of zidovudine.

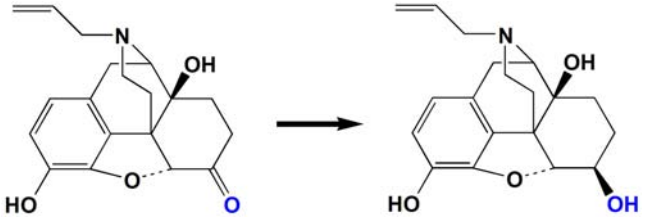


FIGURE 11.11 Chemical equation of reductive reaction of naloxone.

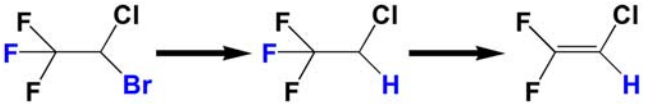


FIGURE 11.12 Chemical equation of reductive dehalogenation of halothane.

(ie, glucose) group from the isoflavone glycoside.[69,70] Only then can the isoflavone aglycone (no sugar group attached) be substantially absorbed into the intestine. Furthermore, hydrolysis by β-glucuronidases derived from bacteria in the colon or enterocytes of the upper small intestine is also a necessary step in the enteric recycling and local recycling of isoflavones. This step hydrolyzes the phase II metabolites of isoflavones back to the isoflavone aglycone form in the intestine, which can be reabsorbed.

11.3.1.3.2 Common hydrolysis substrates

Three main groups of compounds subjected to hydrolysis are amides, carbamates/hydrazides, and esters. Some important compounds metabolized in this way are aspirin (Fig. 11.13), lidocaine (Fig. 11.14), and isoniazid (Fig. 11.15). Hydrolysis of esters usually takes place through both nonspecific and specific esterases. Amides can also be hydrolyzed via nonspecific esterases in the plasma. However, amidases in the liver are more specific and have shown greater amide-hydrolyzing capacity. Hydrazide and carbamate

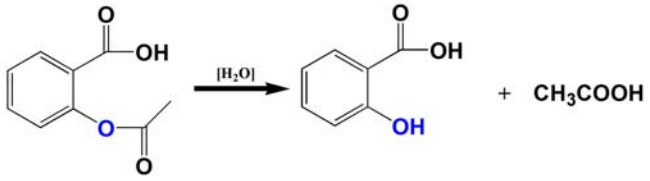


FIGURE 11.13 Chemical equation of ester hydrolysis, using aspirin as an example.

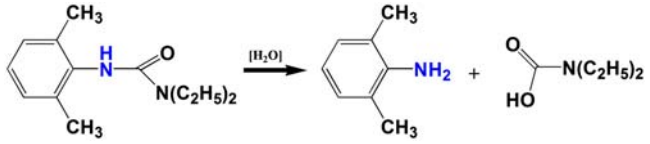


FIGURE 11.14 Chemical equation of amidehydrolysis, using lidocaine as an example.

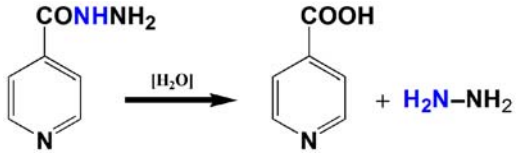


FIGURE 11.15 Chemical equation of hydrazide hydrolysis, using isoniazid as an example.

groups are the least common functional groups subjected to hydrolysis.

11.3.2 Phase II metabolism

Phase II enzymes catalyze conjugating reactions in the classification of drug metabolic pathways. Although drugs can undergo phase I and phase II reactions simultaneously, one pathway typically dominates. For most prescription drugs and drug candidates under development, the predominant pathway is not phase II metabolism. Instead, phase II metabolism often occurs following phase I metabolism, which serves to eliminate drug molecules from the body.

The main metabolic pathways of some compounds involve phase II conjugations. For example, the major metabolic pathway of polyphenol is phase II conjugation. Once absorbed, polyphenols are subjected to three types of conjugation (namely, glucuronidation, sulfation, and methylation).[2,71–73] Glucuronidation and sulfation metabolize polyphenols into very hydrophilic conjugates, whereas methylation produces metabolites that are similarly (if not slightly more) lipophilic than the substrates.[2] Upon methylation, the products (which are still phenols) are usually conjugated subsequently into glucuronides and sulfates since methylation primarily occur on catechols. Therefore, conjugation enzymes that catalyze the formation of hydrophilic metabolites significantly affect the bioavailability of drugs or xenobiotics.

11.3.2.1 *UDP-glucuronosyltransferases or UGTs*

Human UDP-glucuronosyltransferase (UGT) enzymes mainly catalyze the reaction referred to as *glucuronidation*, which involves the covalent linkage of glucuronic acid, derived from the cofactor UDP-glucuronic acid (UDP-GA), to a substrate bearing a suitable functional group (Fig. 11.16). Nucleotide sequences encoding 22 human UGT proteins have been identified, and these sequences have been classified into four gene families, including UGT1 on human chromosome 2q37, UGT2 on chromosome 4q13, UGT3 on chromosome 5p13.2, and UGT 8 on chromosome 4q26. The UGT1 and UGT2 families are the most important in terms of drug metabolism, with the highest efficiency obtained using UDP-GA as the glycosyl donor. By contrast, UGT3 and UGT8 enzymes are generally inactive with UDP-GA as the sugar donor, instead, they utilize UDPGlc, UDPGlcNAc, or UDPGal. Evidence suggests that the enzymes of these two families are unlikely to contribute significantly to drug metabolism.[74]

At the cellular level, all UGTs are membrane-bound proteins located in the smooth endoplasmic reticulum of many tissues.[75] The liver has been well established to contain the most diverse and abundant UGT enzymes. With the exception of UGT1A5, 1A7, 1A8, 1A10, and 2A1, all members of the UGT1 and UGT2 families are expressed in the human liver. In addition to hepatic expression, UGT1A and 2B enzymes are differentially expressed in other tissues, including the kidneys and GI tract, which are considered the most important sites of extrahepatic drug metabolism. The general patterns of UGT expression in the small intestine and colon closely resemble that in the liver, each with a large number of significantly expressed enzymes that sometimes overlap. However, the expression of UGT2B4, the most abundant isoform in the liver, is limited throughout the GI tract; UGT1A7, 1A8, and 1A10, which are essentially absent from the human liver, are expressed in the small intestine and colon.[74] Significant interindividual variability is notably present in UGT tissue expression.

11.3.2.2 *Sulfotransferases or SULTs*

Sulfotransferases (SULTs) are soluble and highly active cytosolic enzymes. They catalyze the transfer of a sulfate moiety from 3′-phosphoadenosine-5′-phosphosulfate to a hydroxyl group on various substrates.[76,77] Compared with UGT isoform-catalyzed metabolism, much less is known about how SULT isoforms metabolize their substrates and the relationships

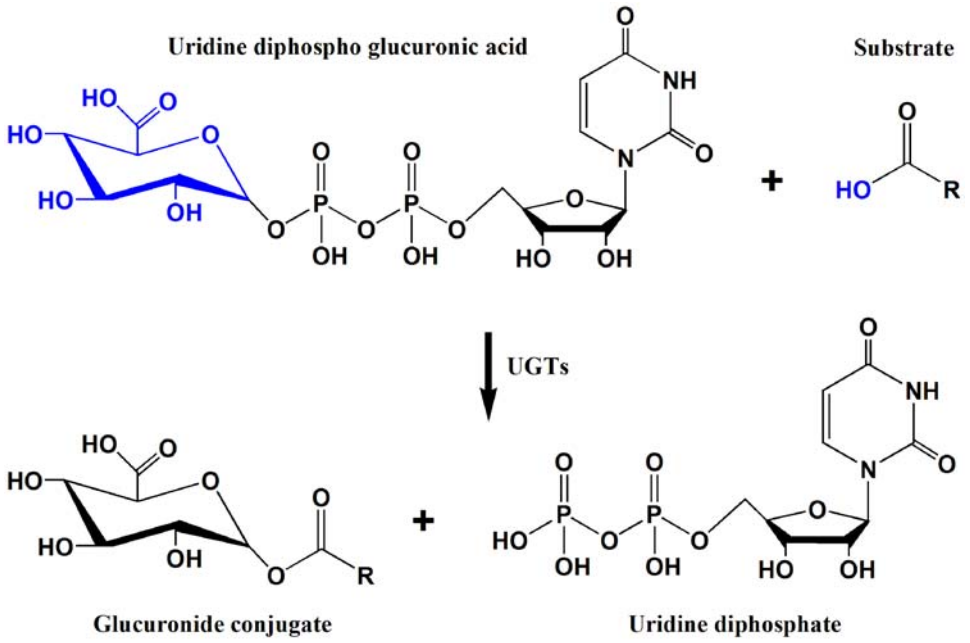


FIGURE 11.16 Glucuronidation reaction scheme.

of structural activities involved. The latter may be due to the following facts: sulfation is often the secondary metabolic pathway (to glucuronidation and other phase II reactions), only a handful of isoforms are commercially available, and commercially available SULTs are much more expensive than UGTs. Therefore, isoforms that are specifically involved in the conjugation of polyphenols have yet to be clearly identified. However, sulfation can still account for a large percentage of the phase II metabolism at lower concentrations and may serve as the primary metabolic pathway for certain compounds such as resveratrol.[78]

11.3.2.3 *Glutathione transferases or GSTs*

Similar to UGTs and SULTs, glutathione transferases (GSTs) are also important in the detoxification process.[79,80] The metabolic products of GSTs are very polar and require the assistance of efflux transporters to exit the cells. More important, glutathione depletion can significantly affect the function of this enzyme. Compared with UGTs and SULTs, GSTs appear to be superior in detoxifying carcinogenic and oxidizing agents. Xenobiotics and natural polyphenols are capable of inducing GSTs, which is critical to maintain an organism's ability to fight oxidative stress and cancer (in mammals).[79,81]

11.3.2.4 *Other conjugating enzymes*

Unlike UGTs and SULTs, methyltransferases such as catechol-*O*-methyltransferase catalyze the *O*-methylation of several catechol-containing polyphenols, thereby forming metabolites that are just as polar as or slightly less polar than the parent compound.[80] This *O*-methylation reaction is subjected to strong inhibitory regulation by *S*-adenosyl-L-homocysteine, which is formed in large quantities during the *O*-methylation of tea polyphenols. Although methylation may not be the major metabolic pathway, its importance in the mechanisms of tea polyphenols cannot be underestimated; methylated metabolites tend to have different activities and some may resist inactivation via UGTs and SULTs.

Similar to methyltransferases, many other phase II enzymes such as acetyltransferases also produce fewer hydrophilic molecules. However, glutamyltransferases and other transferases produce more hydrophilic metabolites to facilitate their elimination.

11.4 PATHWAYS OF DRUG ELIMINATION

The absorptive transport of foods, drugs, or xenobiotics across the intestinal lumen is the first step for them to become bioavailable. As stated previously, the three general routes of absorptive transport include passive transport, facilitated diffusion, and active transport. The actual transport route of these molecules typically depends on their sizes and respective chemical polarities. In general, small hydrophobic molecules are passively transported across the lumen, whereas many hydrophilic nutrients are transported via protein transporters (as discussed earlier in this

chapter). Traditionally, research has focused on how to facilitate the absorptive transport of drugs.

More recently, more attention has been paid to active transport, by which molecules can be actively effluxed out of the intestinal lumen. In addition to excreting or eliminating absorbed molecules, these transporters may be responsible for eliminating phase I and/or phase II metabolites of drugs (as described in detail in the earlier sections of this chapter). To date, several processes of active efflux have been identified. Active transport proteins, such as P-gps, MRPs, and other OATs s, play a major role in governing the intestinal absorption and metabolism of drugs, together with the action of hepatic enzymes and efflux transporters, which determine the bioavailability of drugs and xenobiotics.

11.4.1 P-Glycoprotein

11.4.1.1 Introduction to P-gp

The history of P-gp goes hand in hand with human cancer and chemotherapy. Upon the discovery of cytotoxic drugs that can destroy cancer cells, researchers also discovered that drug resistance (for multiple drugs) can impede the effects of these drugs. P-gp was the first active transporter protein to be discovered, over 30 years ago. Two research groups greatly contributed to the discovery of P-gp and determined this transporter as the major culprit responsible for multidrug resistance (MDR). In 1968, the Ling Lab at the University of Toronto initially used somatic cell genetics to investigate colchicine resistance in Chinese hamster ovary cells, which linked P-gp to MDR.[82] In the meantime, Dr Michael Gottesman's lab at the National Cancer Institute was able to clone the amplified DNA sequences related to the MDR phenotype in cell lines.[83] From both of their works, we now know that MDR and P-gp genes are identical. Today, we know that P-gp and other more recently identified proteins act as efflux transporter pumps with a broad specificity for a variety of substrates.

11.4.1.2 Structure of P-gp

P-gp is a member of a diverse family (over 50 members) of ATP-binding cassette (ABC) transporters that efflux xenobiotics and certain endogenous substances such as bile acids and bilirubin. Dependent on ATP, the P-gp structure shares extensive sequence homology and domain organization compared with other members of this superfamily. The most conserved region is the ABC domain range, approximately between 200 and 250 amino acids consisting of two nucleotide-binding domains (NBDs) located toward the cytoplasmic face of the membrane.[84] More specifically, P-gp contains 12 transmembrane regions split into two halves, each of which contains an NBD.

11.4.1.3 Nomenclature of ABC transporters

P-gp and other members of the MDR-type transporters are organized and classified under the MDR/TAP (transporter associated with antigen processing) protein subfamily. Within the ABC superfamily, MDR-type transporters, such as P-gp (also known as MDR1), MDR3, and BSEP (bile salt export pump, also known as the sister of P-gp), are grouped within the B subfamily and are designated as ABCB1, ABCB4, and ABCB11, respectively.

11.4.1.4 Substrates for P-gp

Substrates for P-gp are typically lipophilic and cationic. As previously mentioned, a large number of compounds are substrates/inhibitors for P-gp, including anticancer agents, antibiotics, antivirals, calcium channel blockers, immunosuppressive agents, and plant chemicals normally found in the human diet.[85]

11.4.1.5 Disruption of P-gp activity

The nature of P-gp is to limit the absorption of compounds by driving the efflux back into the lumen, which in turn limits the bioavailability of these compounds. Data from both in vitro and in vivo studies use human intestinal epithelial cell lines[86] and P-gp-knockout mice,[87] respectively, showed that disruption of P-gp activity could lead to potentially hazardous consequence with regard to drug disposition. Essentially, disruption of P-gp means elimination of a compound in the body via excretion will be substantially decreased (if the compound is predominantly eliminated by P-gp), thereby increasing sensitivity to toxicity as a result of a net increase in compound exposure.

11.4.2 Multidrug-resistance associated proteins

11.4.2.1 Introduction to MRPs

MRPs were discovered much later compared with P-gps. For instance, MRP1 was first cloned in 1992 from a human small cell lung cancer cell line that did not express P-gp.[88] MRP2 was later cloned in 1996 from rat liver and designated as cMRP,[89] as well as from human liver, and designated as a canalicular multispecific OATs.[90] MRP3 was discovered later and was found to be most closely related to MRP1, sharing 58% amino acid identity.[91] In general, all currently known MRPs (MRP1-9) were discovered later than P-gp, and relatively little information exists regarding the function and expression patterns of the more recently discovered MRPs (MRP7-9).

11.4.2.2 Structure of MRPs

As mentioned in the section entitled "Structure of P-gp," earlier in this chapter, the general structure of P-gp and MRPs comprises 12 transmembrane regions split into two halves, each of which contains an NBD. MRP structures are only slightly more complex than P-gp. MRP1-3 and MRP6-7 have a total of 17 transmembrane regions, with another 5 transmembrane regions at the N-terminus. However, these 5 regions do not have an additional NBD. More interestingly, MRP4 and MRP5 do not have these additional regions and thus resemble that of P-gp, which has only 12 transmembrane regions.

11.4.2.3 Nomenclature of MRPs

The nomenclature of MRPs is consistently updated because more novel MRP transporters are being discovered. The subfamily that incorporates MRPs is called the MRP/CFTR (cystic fibrosis transmembrane conductance regulator), which includes MRP1−MRP9, CFTR, and sulfonylurea receptors 1 and 2.

11.4.2.4 Substrates for MRPs

The substrates of MRPs differ among each individual isoform and their localization. However, the overall mechanistic function of MRP is to eliminate compounds from the cell through efflux. For MRPs, studies have shown that this transporter can confer resistance to cytotoxic drugs such as vincristine,[92] peptides,[93] heavy metal anions,[94] and endogenous metabolites such as bilirubin glucuronides.[95] More specifically, MRP2 is centrally involved in the detoxification process through its ability to secrete metabolites into bile. The importance of MRP2 is clearly seen in patients who suffer from Dubin-Johnson syndrome, with defective MRP2 that impairs the ability to secrete glutathione and its conjugates, as well as bilirubin glucuronides.[96] MRP3 exhibits affinity for bile salts and is especially important in mediating bile salt reabsorption. While it is clear that some major overlaps exist with regard to substrate specificity for these MRPs, they do seem to have specific roles in the disposition of xenobiotics and endogenous compounds. Much has yet to be discovered with regard to the substrate specificities and functions of the recently discovered MRP transporters (MRP7-9).

11.4.2.5 Disruption of MRP activity

As previously mentioned, the substrate specificity of MRPs can differ, as can their localization in the cell. Therefore, the effect of the disruption of MRP activity on the disposition process of xenobiotics or endogenous compounds can differ greatly. In general, various MRPs seem to overlap with regard to substrate specificity. Thus, disruption of a specific MRP isoform, such as MRP3, will probably not result in major physiological dysfunction because MRP1 (if expressed and functional) functions similarly to MRP3. However, this is not the case for all MRPs, as certain isoforms (ie, MRP2) seem to be more critically important to the body than others. As mentioned in the previous section, dysfunction of MRP2 in humans can cause Dubin-Johnson syndrome.[96] Data from our laboratories also illustrate the vital role of MRP2 in the efflux of metabolite conjugates with respect to the enterohepatic-recycling process of isoflavones and flavonoids.[101−104]

11.4.3 Breast cancer resistance protein

11.4.3.1 Introduction to BCRP

Human BCRP is one of three major ABC proteins involved in drug disposition. It is encoded by the *ABCG2* gene, which is located on chromosome 4q22. The gene was named as such because it was initially cloned from an MDR breast cancer cell line, where it was found to confer resistance to chemotherapeutic agents. BCRP is physiologically expressed in a variety of tissues, most abundantly in the GI tract, liver, kidney, brain endothelium, mammary tissue, testis, and placenta. It has a significant role in limiting oral bioavailability and transport across the blood−brain barrier, blood-testis barrier, and the maternal-fetal barrier of certain drugs and substrates.[105,106] The role of BCRP in drug disposition can now be fully appreciated because of its wide distribution, expression, and broad substrate and inhibitor specificity. BCRP is recognized by the FDA to be one of the key drug transporters involved in the disposition of clinically relevant drugs.

11.4.3.2 Structure of BCRP

BCRP is a "half ABC transporter" consisting of 655 amino acids and 6 TMDs.[107] It is a member of the subfamily G of the human ABC transporter superfamily, which, in humans, consists of five homologous half transporters, with the ABC domains located toward the N-terminus of the polypeptide chain. BCRP contains only one NBD and one MSD; these features distinguish the protein from P-gp and MRP1, which contain two tandem-repeated halves. In addition, the NBD in BCRP precedes its MSD, a domain organization contrasting that of P-gp and MRP1.[108,109]

11.4.3.3 Nomenclature of BCRP

BCRP is the second member of the subfamily G of the human ABC transporter superfamily; hence, it is also named ABCG2. BCRP was also independently cloned by two other groups named ABCP (for the

human placenta–specific ABC gene)[110] and MXR.[111] MXR is cloned from mitoxantrone-resistant S1-M1-80 human colon carcinoma cells.

11.4.3.4 Substrates for BCRP

ABCG2 is a drug transporter with a broad range of endogenous and exogenous substrates, including both positively and negatively charged and large amphiphilic molecules. Substrates of BCRP were initially reported to include a wide range of chemotherapeutics, such as mitoxantrone, camptothecin derivatives, flavopiridol, and methotrexate.[108] In addition to chemotherapeutics, drugs including prazosin, glyburide, cimetidine, sulfasalazine, and rosuvastatin have been shown to be BCRP substrates.[108,112] Recently developed tyrosine kinase inhibitors such as imatinib, gefitinib, and nilotinib are also known BCRP substrates.[108] A variety of photosensitizers, including pheophorbide A, protoporphyrin IX, and related compounds, are also BCRP substrates.[113] Besides transporting neutral substrates, BCRP transports conjugated organic anions, particularly sulfated and glucuronide conjugates and organic conjugates of drugs, xenobiotics, and endogenous substances.[108] Other BCRP substrates include chemical toxicants, phototoxic compounds, and flavonoids.[108,112] Collectively, BCRP has a broad range of substantially overlapping substrate specificity, which is distinct from that of P-gp or MRP1.

11.4.3.5 Disruption of BCRP activity

As previously mentioned, BCRP actively excretes a wide variety of chemically unrelated compounds from the cells. This protein protects mammalian cells and tissues against various xenobiotics and has a crucial role in the intestine, liver, placenta, and the blood–brain barrier. Disruption of BCRP activity can differ greatly with respect to drug toxicity or the prevention of drug toxicity. Moreover, the role of BCRP in drug disposition is now fully appreciated because it resembles P-gp in tissue distribution and expression. Similar to P-gp, BCRP has a broad substrate specificity. Thus, disruption of BCRP activity could lead to potentially hazardous consequences with regard to drug disposition (ie, drug-drug interactions).

11.4.4 Organic anion transporters

11.4.4.1 Introduction to OATs

The discovery of OATs began almost 70 years ago. At that time, the anionic dye phenolsulphophthalein was observed in very high concentrations in the renal convoluted tubules, which suggested a tubular secretion process. This finding essentially gave rise to the discovery of organic anion (and cation) transport.

11.4.4.2 Structure of OATs

Organic anion transporter 1 (OAT1) was first cloned from rat kidney in 1997 and functionally identified as an OAT.[114,115] OATs and OCTs share a predicted 12-TMD structure with a large extracellular loop between TMD 1 and 2, carrying potential *N*-glycosylation sites. Conserved amino acid motifs exhibited a relationship with the sugar transporter family within the major facilitator superfamily.

11.4.4.3 Nomenclature of OATs

Although the type I and type II classifications of OCTs are formally recognized, this is not the case with OATs. Currently, the most useful method is to acknowledge that OATs can be separated and classified with respect to molecular weight, net charge, and hydrophobicity. Therefore, type I OATs are relatively small (generally <400 kDa) and associated with monovalent compounds such as *p*-aminohippurate. Type II OATs are bulkier (generally >500 kDa), polyvalent in nature, and associated with compounds such as calcein.

11.4.4.4 Substrates of OATs

Substrates for OATs typically include weak acids that have a net negative charge on carboxylate or sulfonyl residues at physiological pH. Although OATs can perform endogenous transport via active secretion of the proximal tubule of compounds, such as riboflavin, the principal function of OATs is to eliminate excess levels of xenobiotic agents.[116] This includes many of the products of phase I and phase II hepatic biotransformation, as well as anionic drugs.[117]

11.4.4.5 Disruption of OAT activity

Disruption or alteration of OAT activity because of underexpression or overexpression is reflected in its pathophysiological states. Essentially, accumulation of uremic toxins (of which many are organic anions) in the kidney as a result of OAT dysfunction will lead to general renal disease.[118]

11.5 COUPLING OF ENZYMES AND EFFLUX TRANSPORTERS

Until recently, conjugating enzymes such as UGTs and their ability to metabolize compounds (phase II metabolism) were believed to be the most important steps in the disposition of xenobiotics. The ability of UGTs to biotransform a lipophilic compound into a more hydrophilic metabolite (via the addition of a glucose sugar molecule) was initially believed to be the major biological barrier to the entry of xenobiotics/

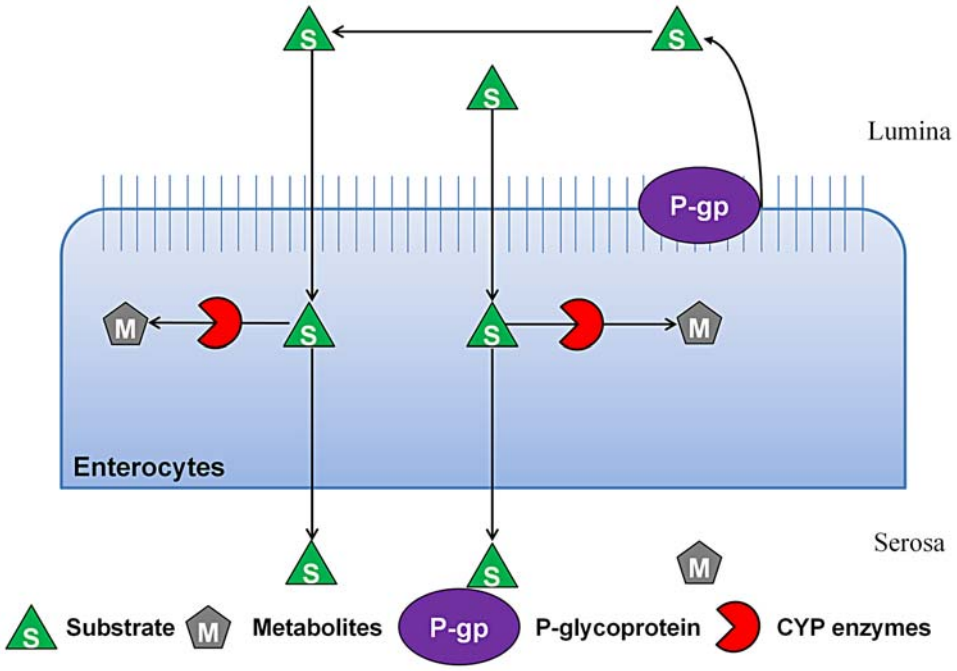


FIGURE 11.17 Schematic representation of the double jeopardy theorem. Substrates are represented by triangles, whereas the products of enzymes or metabolites are represented by pentagons or hexagons.

drugs into systemic circulation, representing one of the main pathways for their elimination.[2] However, this belief places great emphasis on UGT activity and does not comprehensively cover the overall disposition of xenobiotic/drugs. The fallacy lies in the fact that many conjugates are hydrophilic and cannot permeate the cell membrane once biotransformed by UGTs. While UGT metabolism is a highly necessary component for the disposition process of xenobiotics/drugs, more focus has been placed on various transporters that are capable of transporting negatively charged hydrophilic conjugates across the cellular membrane. In this section, we discuss three theories that explain the mechanistic details and complexities of this coupling process.

11.5.1 Double Jeopardy theorem

The double jeopardy theorem, in a slightly different form, was originally proposed to explain the mechanistic details that describe the coupling phenomenon between a phase I metabolic enzyme, such as CYP3A4, and an efflux transporter, such as P-gp (Fig. 11.17). The theorem was originally intended by Benet and coworkers to explain the pharmacokinetic behavior of certain drugs that are substrates of both CYP3A and P-gp.[115,116]

11.5.1.1 Mechanistic description of the theorem

The mechanism of action for this coupling theory is slightly less complex than the revolving door theorem (discussed next). In essence, the substrate is assumed to be absorbed (at the apical side of the membrane) and metabolized by the enzyme CYP (typically CYP3A4). Afterward, P-gp can take this one-time metabolized substrate and transport it via efflux back into the intestinal lumen. It is important to note that this one-time metabolized substrate can pass through the other side (basolateral) of the membrane and thus reach the systemic circulation. However, if this one-time metabolized substrate is effluxed back into the intestinal lumen via P-gp, it can be absorbed one more time at the apical side of the membrane (further down the intestinal tract), metabolized again by the same enzyme CYP (albeit in a different cell), and subjected to the same metabolism. Owing to these two repeats of "prosecution," the theory merits the term "double jeopardy."[121,122]

11.5.1.2 Consequences of disruption

In general, the consequences of the disruption of this mechanism and any other coupling mechanism can vary depending on the substrate in question. Given differences between substrate disposition profiles, the toxicity and interactions of certain drugs with other drug mechanisms, as well as the compensatory mechanism or effects by other enzymes or transporters in the system, will vary accordingly. In terms of the double jeopardy theorem (and, more specifically, CYP and P-gp), disruption of P-gp activity would result in

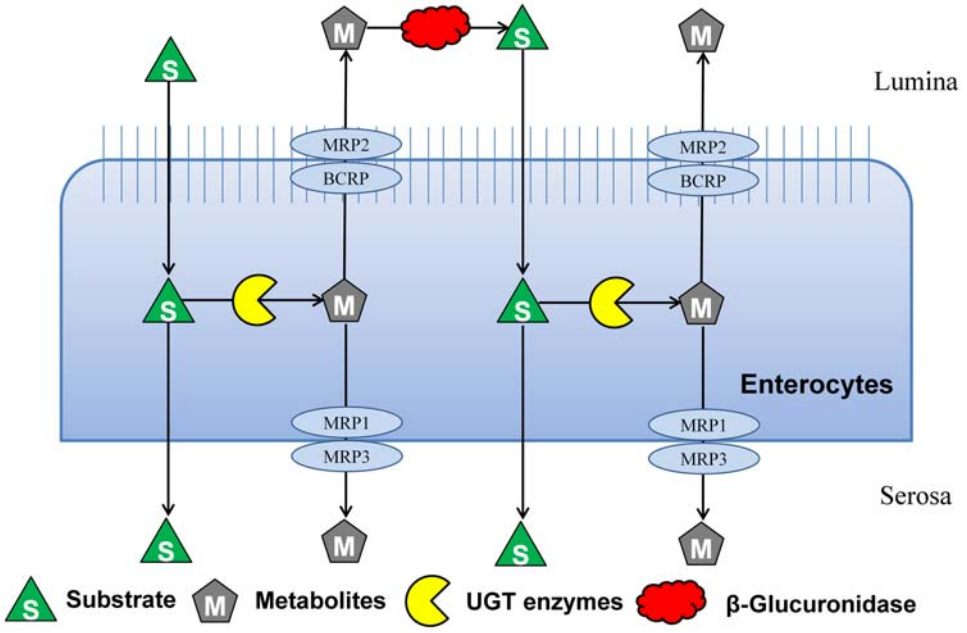


FIGURE 11.18 Schematic representation of the revolving door theorem. Substrates are represented by triangles, whereas products of enzymes or metabolites are represented by pentagons.

less exposure time of the substrate in question at the apical membrane. This shorter exposure time leads to higher substrate concentrations inside the intestine (ie, increase in absorption) and could potentially decrease the bioavailability of the substrate, its metabolite in the systemic circulation, or both. On the other hand, if CYP activity was disrupted, the amount of metabolite available would decrease severely, allowing more substrates to reach the systemic circulation.

The revolving door theorem, recently proposed, hypothesizes that efficient coupling of conjugating enzymes (UGTs) and efflux transporters enables enteric recycling processes.[2] Therefore, unlike the previous theory, the revolving door theorem emphasizes the importance of efflux transporters and their ability to solely enable the cellular excretion of hydrophilic metabolites (Fig. 11.18). This theory also recognizes the transport of hydrophilic metabolites to both apical and basolateral sides of the membrane. This latter point adds complexity to the theory and its mechanism (discussed next).

11.5.2 Revolving door theorem

11.5.2.1 Mechanistic description of the theorem

The mechanism of the revolving door theorem is best illustrated step by step for a model compound, such as the soy isoflavone, genistein (aglycone). When absorbed into the intestinal lumen, genistein undergoes two possible immediate fates. First, genistein passes through the basolateral side of the membrane and reaches the systemic circulation. However, research has found that this fate is not always the case, as bioavailability is extremely low for this compound upon ingestion.[123] Second, in the intestinal lumen, genistein is likely to be extensively metabolized by UGTs.[123,124] After biotransformation by UGTs, the genistein metabolite again has two fates. One, as an organic anion, the genistein metabolite can be transported by efflux transporters (MRP1 and/or MRP3) located at the basolateral side of the lumen and eventually reach the systemic circulation. Second, the metabolite can be transported back across the apical side of the intestinal lumen by efflux transporters, such as MRP2 (or perhaps BCRP). According to our research,[125] this efflux step is most likely the rate-limiting step of this coupling process. In other words, the amount of metabolites produced by UGTs may not significantly alter or influence the disposition of genistein or its metabolites because of the dependence on the efflux step by transporters. This dependence is ultimately the theme of the revolving door theorem, which places emphasis on the rate of the efflux step (ie, the speed at which the revolving door is spinning) as the determining factor for the disposition of genistein and its metabolite.

Upon efflux into the intestinal lumen, genistein metabolites are converted back to the genistein aglycone through bacterial hydrolysis and can thus reenter the intestinal lumen. This entire process, which occurs at the intestinal level, is known as *enteric recycling*. This

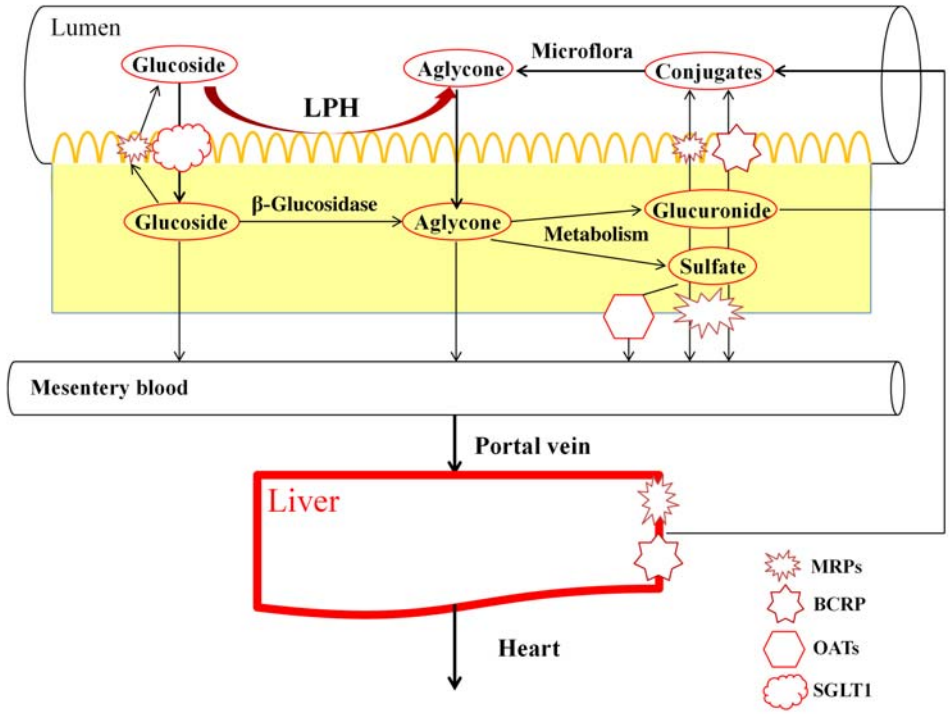


FIGURE 11.19 Schematic representation of both enteric and enterohepatic recycling with that isoflavone aglycone as an example. Isoflavones are typically subjected to rapid hydrolysis with intestinal gut hydrolase. Absorption of aglycone occurs rapidly, and the isoflavone is subjected to UGT metabolism with metabolites that typically efflux at both the intestinal and hepatic levels. Further hydrolysis from conjugates back to the aglycone form completes the cyclical recycling process, which features repeated absorption, followed by metabolism and eventual efflux.

extensive metabolism at the enteric level is also known as *first-pass metabolism*. Regarding the liver, genistein passing UGTs in the intestine is subjected to liver UGTs. These metabolites, as well as any metabolite that passes through the intestinal level, will again be secreted in bile into the intestinal lumen, hydrolyzed, and become ready for absorption. This process is known as *enterohepatic recycling*; once again, the revolving door theorem is considerably involved (Fig. 11.19).

11.5.2.2 Consequences of disruption

The consequences of disruption of this coupling theory depend on the substrate and its metabolites in question, as the body possesses an elaborate system of various metabolic enzymes (ie, UGTs and SULTs) and efflux transporters (ie, MRPs, BCRP, and OATs) with overlapping substrate specificities. Considering that detoxification is an essential function for organism survival, we have developed a highly complex system through evolution featuring multiple enzymes and efflux transporters that can compensate if any particular member becomes dysfunctional. Therefore, a wide number and variety of enzymes (ie, both UGTs and SULTs) must be made dysfunctional to disrupt coupling severely.

11.5.3 Enteric and enterohepatic recycling

Besides metabolism, various recycling or circulation processes are known to affect the oral bioavailability of certain phenolics (including flavonoids) and their conjugates. In addition to enteric recycling, excretion could occur in the liver when mediated by efflux transporters. In the liver, metabolites could be excreted into the intestinal lumen via bile. After hydrolysis of the metabolites to substrates, the substrates could then be reabsorbed. Enterohepatic recycling restricts the entry of substrates into systemic circulation but extends the systemic exposure time, which could explain the poor systemic bioavailability but reasonably long half-lives of certain phenolics that participate in recycling (Fig. 11.20).[126]

The mechanism of enterohepatic recycling can be best illustrated by using flavonoids as model compounds. The phase II metabolites of flavonoids formed in hepatocytes are hydrophilic. Therefore, flavonoids have difficulty penetrating the cell membrane by passive diffusion. Like enterocytes, hepatocytes possess basolateral and apical sides.[127] The expression of transporters on the basolateral and apical sides of hepatocyte membranes is obviously different. MRP3, MRP4, MRP5, and MRP6 are located on the basolateral side, whereas P-gp, BCRP, and MRP2 are located on the apical side.[128] In general, details of the

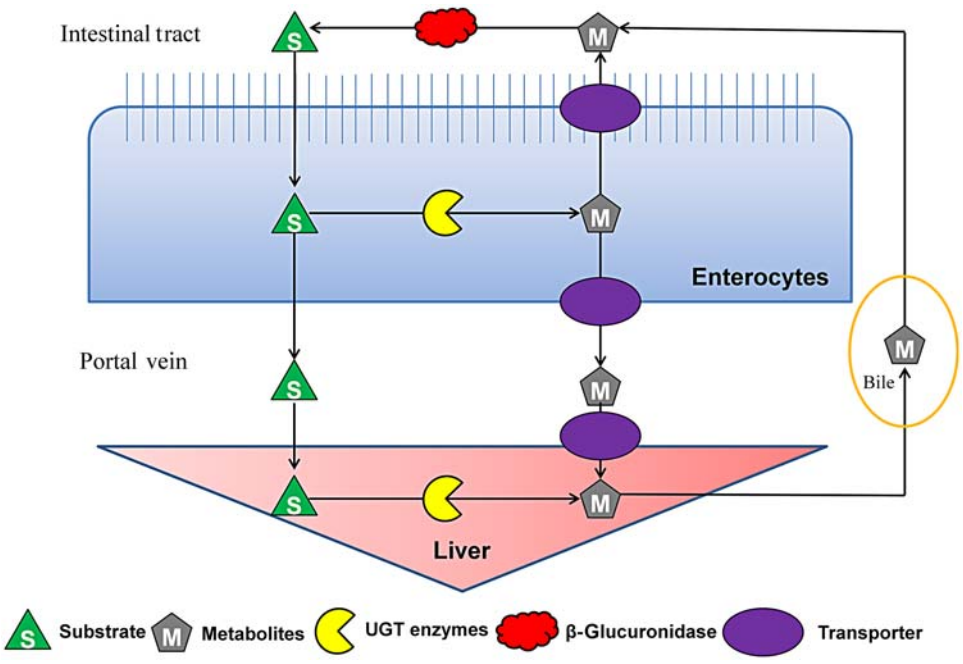


FIGURE 11.20 Schematic representation of enteric and enterohepatic recycling. Substrates are represented by triangles, whereas the products of metabolism or metabolites are represented by pentagons.

efflux of flavonoid metabolites by transporters are revealed less in hepatocytes. However, as the same transporter is expressed in both hepatocytes and enterocytes, we can assume that the same transporters perform these functions. Phase II metabolites of flavonoids are commonly excreted via bile. After duodenal administration, high concentrations of flavonoid metabolites could be detected in the bile.[129] In the rat intestine perfusion model, considerable amounts of metabolites could be recovered from the bile.[126,130,131] Enterohepatic recycling of phase II conjugates of flavonoids is common. Glucuronides and sulfates in flavonoids could be converted back to aglycones by hydrolases and reabsorbed in the intestinal tract, completing a process called *enterohepatic circulation*. Enterohepatic recycling contributes in part to the prolonged presence of relevant radioactivity in plasma, which may enhance the efficacy of flavonoid metabolites.[132]

11.6 REGULATION OF TRANSPORTERS AND ENZYMES BY NUCLEAR RECEPTORS

11.6.1 Nuclear receptors

Nuclear receptors (NRs) are ligand-inducible transcription factors that specifically regulate the expression of target genes involved in metabolism, development, and reproduction.[133] NRs include the pregnane X receptor (PXR; NR1I2), constitutive androstane receptor (CAR; NR1I3), liver X receptors α and β (LXR; NR1H3 and NR1H2), farnesoid X receptor (FXR; NR1H4), and peroxisome proliferator-activated receptors α, β/δ, and γ (PPAR; NR1C1, NR1C2, and NR1C3). Most of the NRs share common structural features, which comprise two essential functional domains that include the N-terminal DNA-binding domain (DBD) and a C-terminal ligand-binding domain (LBD). The conserved DBD encompasses two DNA-binding zinc-finger motifs that bind to distinct target genes. The LBD, which is connected to the DBD through a hinge region, folds to form a hydrophobic pocket, into which the ligand binds.[134] In the absence of xenobiotics, the basal expression level of proteins mediated by NRs is low. On ligand binding, the NRs transcriptionally regulate the expression of target genes involved in the biotransformation and homeostasis of both exogenous and endogenous substances by the recruitment of coactivators or corepressors (Fig. 11.21).

11.6.2 Pregnane X receptor and constitutive androstane receptor

PXR and CAR both belong to the NR superfamily and are well-established xenobiotic sensors that are capable of binding to various structurally diverse chemicals. They are predominantly expressed in the liver, and both are involved in the regulation of drug metabolism, bile acid transport, cancer, and cholesterol metabolism, among others. When bound to and activated by ligands,

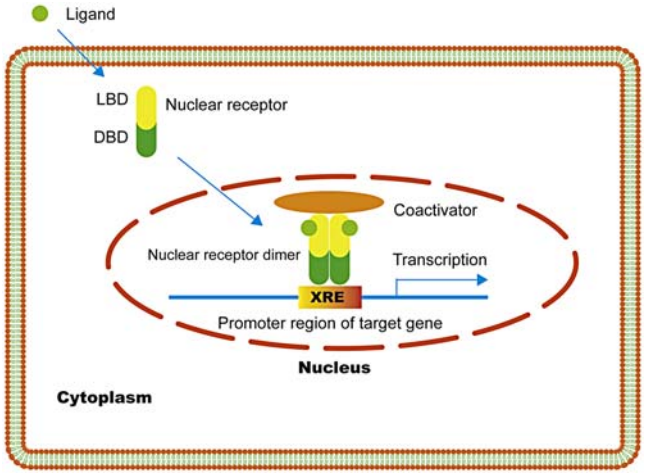


FIGURE 11.21 Model of nuclear receptor signaling. *Source: Derived from Sugatani J, Kojima H, Ueda A, Kakizaki S, Yoshinari K, Gong Q.H, et al. The phenobarbital response enhancer module in the human bilirubin UDP-glucuronosyltransferase UGT1A1 gene and regulation by the nuclear receptor CAR.* Hepatology (Baltimore, MD) *2001;33:1232–8.*[135]

the sensors translocate from the cytoplasm to the nucleus of cells, where they bind to DNA response elements as a heterodimer or heterotetramer with the retinoid X receptor (RXR).[136] Upon ligand binding, PXR and CAR are activated to control the expression of a large number of target proteins involved in the biotransformation and homeostasis of both exogenous and endogenous substances, including DMEs and transporters, thereby influencing the detoxification of xenobiotics.

11.6.3 Regulation of transporters and enzymes by PXR

PXR is one of the NRs that are originally identified as xenobiotic receptors that regulate CYP gene expression. CYP3A4, the most highly expressed CYP enzyme in human liver, is highly inducible by a wide variety of xenobiotics. CYP3A induction by xenobiotics is now believed to be largely due to xenobiotic binding and activation of PXR. Originally, PXR/RXR heterodimers were observed to bind efficiently to DR3-type xenobiotic response elements (XREs) present in the proximal promoters of rat *CYP3A23* and *CYP3A2* genes.[137] PXR/RXR heterodimers were later found to bind to a DR3 element in the human CYP3A4 gene enhancer and an ER6 element located in the proximal promoter of the *CYP3A4* gene.[138,139] In addition to CYP3A family members, PXR regulates other phase I enzyme genes, including *CYP2B6*, *CYP2B9*, *CYP2C8*, *CYP2C9*, and *CYP2C19*. Phase II enzyme genes regulated by PXR include those in the GST, SULT, UGT, and carboxylesterase families. The expressions of certain transporters, including OATP1A4, OATP2 (OATP1B1), MRP2, and MDR1, are also influenced by PXR (Table 11.2).

TABLE 11.2 PXR Target Genes Involved in Phase I, II, and III Metabolism

Phase	Target genes	Species
Phase I	*Aldh1a1*	Mouse
	Aldh1a7	Mouse
	CYP1A1	Mouse, human
	CYP1A2	Human
	CYP1A6	Human
	CYP2B1	Rat
	CYP2B2	Rat
	CYP2B6	Human
	CYP2b10	Mouse
	CYP2C8	Human
	CYP2C9	Human
	CYP2C19	Human
	CYP3A1	Rat
	CYP3A2	Rat
Phase II	*CYP3A4*	Human
	CYP3a11	Mouse
	CYP3A23	Rat
	CYP3a44	Mouse
	CYP7A1	Human
	CYP11A1	Human
	CYP11B1	Human
Phase III	*UGT1A1*	Mouse, human
	UGT1A3	Human
	UGT1A4	Human
	UGT1A6	Human
	UGT2b5	Mouse
	GSTA1	Mouse, rat
	GSTA2	Mouse, rat
	Gstm1	Mouse, rat
	Gstm2	Mouse
	SULT2A1	Mouse, human
Phase IV	*MDR1*	Mouse, human
	MRP2	Mouse, rat, human
	MRP3	Mouse, human
	OATP1A4	Mouse, rat
	OATP2 (OATP1B1)	Mouse, rat

Derived from Sugatani J, Kojima H, Ueda A, Kakizaki S, Yoshinari K, Gong Q.H, et al. The phenobarbital response enhancer module in the human bilirubin UDP-glucuronosyltransferase UGT1A1 gene and regulation by the nuclear receptor CAR. Hepatology (Baltimore, MD) *2001;33:1232–8.*

11.6.4 Regulation of transporters and enzymes by CAR

Similar to PXR, CAR is another NR that was originally identified as a xenobiotic receptor that regulates CYP gene expression. CAR regulates gene expression through the phenobarbital responsive element module (PBREM), in the promoter regions of target genes. In humans, *CYP2B6* is responsible for the metabolism of a large number of drugs. *CYP2B6* is highly inducible not only by phenobarbital (PB), but also by a large number of other compounds.[140] CAR response elements have been shown to be located in human *CYP2B6* and mouse *Cyp2b10* gene promoters. CAR-/- mice fail to display *Cyp2b10* induction in response to PB treatment.[134] In addition to *CYP2B6*, hCAR can regulate the expression of other CYPs, including *CYP2B6*, *CYP2C9*, *CYP2C19*, *CYP3A4*, and *CYP3A5*. UGT1A1 is an important phase II enzyme, the expression of which has been reported to be dually regulated by hCAR. UGT1A1 was initially identified as a CAR target gene when CAR is activated by PB.[141] GST, another phase II conjugation enzyme, is modestly induced by PB[141] and TCPOBOP,[142] thus suggesting that CAR may regulate the expression of GSTs. Besides its regulation of the phase I and II DMEs, CAR regulates the expression of drug transporters, such as P-gp, OATP2 (OATP1B1), and MRP2-4 in mouse liver (Table 11.3).

TABLE 11.3 CAR Target Genes Involved in Phase I, II, and III Metabolism

Phase	Target genes	Species
Phase I	*Aldh1a1*	Mouse
	Aldh1a7	Mouse
	CYP1A1	Mouse
	CYP2a4	Mouse
	CYP2A6	Human
	CYP2B1	Rat
	CYP2B2	Rat
	CYP2B6	Human
	CYP2b10	Mouse
	CYP2C6	Rat
	CYP2C7	Rat
	CYP2C9	Human
	CYP2C19	Human
	CYP3A1	Rat
	CYP3A4	Human
	CYP3a11	Mouse
Phase II	*UGT1A1*	Mouse, human
	UGT2B1	Rat
	GSTA1	Mouse, rat
	GSTA2	Mouse, rat
	GSTA3	Mouse, rat
	Gstm1	Mouse, rat
	Gstm2	Mouse
Phase III	*MDR1*	Mouse, human
	Mrp1	Mouse
	MRP2	Mouse, rat, human
	MRP3	Mouse, human
	Mrp4	Mouse
	OATP2 (OATP1B1)	Mouse, rat

Derived from Sugatani J, Kojima H, Ueda A, Kakizaki S, Yoshinari K, Gong Q.H, et al. The phenobarbital response enhancer module in the human bilirubin UDP-glucuronosyltransferase UGT1A1 gene and regulation by the nuclear receptor CAR. Hepatology (Baltimore, MD) *2001;33:1232–8.*

11.6.5 Regulation of transporters and enzymes by other NRs

In addition to PXR and CAR, other NR members perform important functions in the regulation of DMEs and drug transporters. For example, aryl hydrocarbon receptor, which shares several significant characteristics with CAR and PXR, mainly enhances the gene expression levels of *CYP1A1*, *CYP1A2*, *CYP1B1*, *UGT1A1*, *UGT1A3*, *UGT1A4*, *UGT1A6*,[143] *MDR1/3*, and *MRP2/3*.[144] Vitamin D receptor is related to the expression of several CYP genes, such as *CYP2C9*, *CYP2B6*, *CYP3A4*,[145] and transporter MDR1.[146] Nuclear factor erythroid 2-related factor 2 (Nrf2) mainly regulates the expression of phase II enzymes (GST),[147] as well as transporters such as MDR1, BCRP, MRP2, and MRP3.[148] Peroxisome proliferator-activated receptors (PPARs) mainly affect phase II enzymes, such as SULTs, UGTs,[143] and transporter BCRP.[149]

11.7 PHYSICOCHEMICAL FACTORS AFFECTING DRUG ABSORPTION

We classify the physicochemical characteristics of a drug molecule or particle into those that are solely related to chemical structures and those that are related to physical shape (eg, crystalline forms). Considering the biological constrains posed by GI barriers, we must consider the following physicochemical properties: lipophilicity, size, charge, solubility, dissolution, and ionization.

11.7.1 Lipophilicity

The lipophilicity of a drug or drug candidate molecules is a major concern in the development of dosage forms because drug molecules must penetrate the lipid bilayer of most cellular membranes, including that of enterocytes. Therefore, drug molecules should be lipophilic to allow good absorption.

The lipophilicity of uncharged molecules can be estimated experimentally via a straightforward method of measuring the partition coefficients in a water/oil (eg, water/octanol) bi-phase system. For molecules that are weak acids or bases, one must consider the pH at which the majority of the species remain uncharged versus the pH at which the majority of the species are charged. Based on the pH partition theorem, absorption is favored when a molecule is uncharged, which works for drug molecules that are absorbed via passive diffusion.

This general trend that favors lipophilicity contradicts the fact that higher lipophilicity favors better absorption, as pH partition theory would have suggested. As an explanation, during transcellular absorption in the enterocytes, drug molecules must first enter through apical membrane, travel through the cytosolic domain, and then exit the basolateral membrane. If a compound binds too tightly to the cell membrane as the result of its extremely high lipophilicity, the compound would be unable to enter the aqueous cytosolic domain or exit from the lipophilic basolateral membrane into the aqueous lamina propria. As mature enterocytes require about 10 days to develop and function for about 4 days after maturation, drugs attached to cell membranes (apical or basolateral) would eventually slough off when the mature cells perish as the result of their natural life cycle. According to Lipinski's Rule of Five, the partition coefficient should be positive but less than 100,000 to 1 (or log PC = 5).[150,151] In another statistical study, the log PC was seen to be the best at log PC < 3.[152]

11.7.2 Size

Passive absorption in the GI tract is severely limited by the size of the penetrating drug molecule probably because of the well-organized and packed structure of the lipid bilayer. When a molecule is excessively large, the potential energy resulting from concentration difference is not sufficiently large to generate the high energy required to markedly disturb the bilayer, which is necessary for a large molecule to insert into it and then leave afterward. Therefore, the size (and perhaps the surface area) of a molecule are major factors limiting absorption via passive diffusion. According to Lipinski's Rule of Five, the optimal molecular weight is less than 500.[150,151] In a similar paper, that molecular weight was extended to 550.[152]At any rate, very few drug molecules with large molecular weights are absorbed to an appreciable extent (> 20%). Cyclosporin A is one of the exceptions; this drug exhibits good bioavailability to exert intended effects, and formulations with improved solubility demonstrate increased absorption.[153]

11.7.3 Charge

The effects of charge on the passive absorption of drugs are well recognized. In general, charged molecules are not as permeable as the corresponding uncharged species when the compound is absorbed via passive diffusion. However, the effect of charge on the absorption of drugs via carrier-mediated transport process is complex. Several transporters favor neutral substrates, whereas others prefer positively charged or negatively charged substrates.

11.7.4 Solubility

Solubility affects the absorption of drugs, as it affects the driving force of drug absorption—specifically, the concentration of drug molecules at the site of absorption. Other chapters in this book are devoted to this topic, so we will refrain from any major discussion of it here.

11.7.5 Dissolution

Dissolution affects the absorption of drugs because this property affects the driving force of drug absorption, as stated previously. However, the term describes the process of solubilization and is represented by the dissolution rate. Drug molecules must possess sufficient dissolution rates, as drug particles are scheduled to be in contact with the absorption area for only a finite length of time because of residence time constraints of the GI tract. Other chapters in this book discuss this property in greater detail.

11.7.6 Ionization (pK_a)

Numerous drugs are either weak acids or bases that exist in both ionized and unionized forms in vivo. The ionized part of the drugs cannot cross lipid membranes, whereas the unionized part of drugs can cross freely. The degree of ionization depends on the pK_a of the drug and the pH of the GI tract. The pK_a is the pH at which the drug is 50% ionized and 50% unionized. The ratio of two forms, which can be calculated by using the Henderson-Hasselbalch equation[154]

(as follows), at a particular site influences the rate of absorption.

For acidic drugs,

$$pK_a - pH = \log[AH/A^-] \quad (11.1)$$

If pH is lower than pK_a, more unionized species will be present.

For basic drugs,

$$pK_a - pH = \log[BH^+/B] \quad (11.2)$$

If pH is higher than pK_a, more unionized species will be present.

The unionized form of weak acids ($pK_a = 3-5$) and the ionized form of weak bases ($pK_a = 8-10$) predominate in the stomach (with a pH of 1). Hence, weak acids are more readily absorbed from the stomach than weak bases. Acidic drugs are present in ionized form in the intestines (with a pH of 7) and are thus less absorbed. However, given the small surface area of the stomach and the contrasting very large surface area of the intestine, even acidic drugs are often more absorbed from the upper part of the intestine. Basic drugs are present in unionized form in the intestines and are absorbed at a considerably more rapid rate. However, the effects of solubility in the intestines must be accounted for, as discussed in the following sections.

11.8 BIOLOGICAL FACTORS AFFECTING DRUG ABSORPTION

The oral route of administration is the most common and popular route of drug delivery. The oral absorption process of a drug from a pharmaceutical dosage form is complex. In general, absorption involves four phenomena: (1) dissolution of the drug from its dosage form; (2) solubility of the drug as a function of its physicochemical characteristics; (3) effective permeability of the drug to the intestinal mucosa; and (4) presystemic metabolism of the drug.[149] Moreover, the oral dosage form must be designed to account for extreme pH ranges, the presence or absence of food, degradative enzymes, varying drug permeability in the different regions of the intestine, and motility of the GI tract. A drug may encounter several barriers to absorption when released from its dosage form and upon dissolution into the GI fluids. The drug must remain in solution and not become bound to food or other materials within the GI tract. The drug must be chemically stable so as to withstand the pH of the GI tract, and it must be resistant to enzymatic degradation in the lumen.

For most drugs, the optimum site for drug absorption after oral administration is the upper portion of the small intestine or duodenum region. The unique anatomy of the small intestine or duodenum provides an immense surface area for drug absorption. The large surface area of the intestine is due to the presence of valvelike folds in the mucous membrane, which contain small projections known as *villi*. These villi contain even smaller projections known as *microvilli*, which form the brush border of epithelium. The duodenal region is also highly perfused with a network of capillaries, which helps maintain a concentration gradient from the intestinal lumen and plasma circulation. Overall, the GI absorption of orally administered drug is determined by the permeability of the GI mucosa, the surface area, the concentration gradient, and the transit rate in the GI tract.

Gastric emptying, especially that of food, the pH of the GI, and luminal enzymes are important factors affecting the plasma concentration profile of orally administered drugs. In addition, the intestinal transit rate also exerts a significant influence on drug absorption, as the rate determines the residence time of the drug in the absorption site.[150] In this chapter, the importance of GI biological factors in determining the absorption kinetics and bioavailability of orally administered drugs is discussed.

11.8.1 Transit time

Once a drug is administered orally, the exact location and environment of the drug product within the GI tract are difficult to determine. GI motility tends to move the drug through the alimentary canal, so that the drug does not stay at the absorption site. The transit time of the drug in the GI tract depends on the physiochemical and pharmacologic properties of the drug, the type of dosage form, and various physiologic factors.[151]

Intestinal propulsive movements determine the intestinal transit rate and, as a consequence, the residence time of a drug in the intestine. Evidently, greater intestinal motility results in a shorter residence time and less time for the drug to be absorbed at the site. Peristaltic activity increases after a meal as a result of the GI reflex, which is initiated by distension of the stomach and results in increased motility and secretion.

GI mixing movements are a result of contractions dividing a given region of the intestine into segments, thereby producing an appearance that is similar to a chain of sausages. These mixing motions tend to improve drug absorption since (1) mixing movements can increase dissolution rates and influence absorption and (2) the mixing process can increase contact between a drug and membrane. Transit through the

small intestine appears to deviate in a variety of ways from movement through the stomach. Once emptied from the stomach, the drug (such as pellets and tablets) moves along the small intestine and reaches the ileocecal valve within approximately 2–4 hours. Transit in the intestine appears to be less dependent on the physical nature of the drug, such as liquid versus solid, and size of solids in comparison with that in the stomach. Furthermore, food appears not to influence intestinal transit as it does gastric emptying. While no gender-related differences appear in intestinal transit time, vegetarians appear to have longer intestinal transit times in comparison with nonvegetarians.[152]

The colonic transit of pharmaceuticals is long, variable, and depends on the dosage form, diet, eating pattern, and disease. Contractile activity in the colon can be divided into two main types: (1) propulsive contractions or mass movements, which are associated with the aboral (ie, away from mouth) movement of contents and (2) segmental or haustral contractions, which serve to mix the luminal content and result in small aboral movements. Segmental contractions are brought about by contraction of the circular muscle and predominate intestinal transit, whereas propulsive contractions, which are due to contractions of the longitudinal muscle, occur only three to four times daily in normal individuals. Colonic transit is thus characterized by short bursts of activity followed by long periods of stasis. Movement is mainly aboral (ie, toward the anus). Colonic transit can vary from 2 to 48 hours. In most individuals, mouth-to-anus transit times are longer than 24 hours.[150]

11.8.2 pH

The pH of GI fluids varies considerably along the length of the GI tract. Gastric fluid is highly acidic, normally exhibiting a pH within the range 1–3.5 in healthy people in the fasted state. Following ingestion of a meal, gastric juice is buffered to a less acidic pH, which depends on meal composition. Typical gastric pH values following a meal are in the range of 3–7. Depending on meal size, the gastric pH returns to lower fasted-state values within 2–3 hours. Thus, only a dosage form ingested with or soon after a meal will encounter these higher pH values, which may affect the chemical stability of a drug or its dissolution, absorption, or a combination thereof.[150]

The pH of GI may influence the absorption of drugs in a variety of ways. If the drug is a weak electrolyte, the pH may influence its chemical stability in the lumen, as well as its dissolution or absorption. Chemical degradation due to pH-dependent hydrolysis can occur in the GI tract. The result of this instability is lower bioavailability, as only a fraction of the administered dose reaches the systemic circulation in the form of intact drug.[150] Despite of this possible change, long-term and sustained change of GI pH to improve the absorption of drugs has never been approved by the FDA.

11.8.3 Food

The presence of food in the GI tract can affect the bioavailability of the drug from an oral drug product. Digested foods contain amino acids, fatty acids, and numerous nutrients that may affect intestinal pH and the solubility of drugs. Several effects of food on the bioavailability of a drug product may include (1) delay in gastric emptying; (2) stimulation of bile flow; (3) a change in the pH of the GI tract; (4) an increase in splanchnic blood flow; (5) a change luminal metabolism of the drug substance; and (6) physical or chemical interaction of the meal with the drug product or drug substance.[153]

Besides the general information given here, interactions between food and drugs other than dilution, binding, precipitation, absorption, chelation, changes in gastric, and urinary pH, among others, must be considered. For instance, grapefruit juice may considerably increase the peroral bioavailability of numerous *CYP3A4* substrates by inhibiting intestinal phase I metabolism by up to 300%; drugs that exert these effects include felodipine, nifedipine, nisoldpine, and nitrendipine. Numerous drugs are transported back from the enterocyte into the GI lumen by the P-gp efflux transporter system. However, grapefruit juice also possesses a higher affinity for P-gp than other drugs. Therefore, efflux may be inhibited and bioavailability increases, for drugs such as cyclosporine, vinca alkaloids, digoxin, and fexofenadine. Grapefruit juice also exhibits affinity for OATP, which would inhibit the influx of drugs transported by this system from enterocytes to the plasma, thus decreasing their bioavailability.[154]

11.8.4 Luminal enzymes

Pepsin is the primary enzyme found in gastric juice. Lipases, amylases, and proteases are secreted from the pancreas into the small intestine in response to food ingestion. These enzymes are responsible for most nutrient digestion. Pepsins and proteases are responsible for the degradation of protein and peptide drugs in the lumen. Other drugs that resemble nutrients, such as nucleotides and fatty acids, may also be susceptible to enzymatic degradation. Lipases may affect the

release of drugs from fat/oil-containing dosage forms. Drugs that are composed of esters can be susceptible to hydrolysis in the lumen. Bacteria, which are mainly localized within the colonic region of the GI tract, secrete enzymes that have been utilized when designing drugs or dosage forms to target the colon.[150]

In summary, this chapter presented a general description of all barriers to the oral bioavailability of drugs. In this chapter, we showed that barriers to oral bioavailability present several layers, including solubility and dissolution rate, absorption, metabolism, efflux, and excretion. We have not discussed the transport of drugs to target organs, which may also serve as an important barrier for drugs to become bioavailable in the CNS, tumor cells, and difficult-to-access organs. Our understanding of the solubility and dissolution behaviors of solids has not changed considerably over the last 20 years in that we still cannot accurately predict solubility in vivo. However, we have made great progress in the last 15 years on research into the in vivo transport of drugs. However, more than 90 human transporters with unknown functions and substrates exist. Therefore, our prediction of drug absorption remains inadequate. The use of Caco-2 in the discovery process is helpful in addressing this bottleneck. The metabolism and excretion of drugs has received considerable attention from regulatory agencies because of drug-interaction issues and the difficulties associated with the prediction of issues, especially when enzymes and transporters may work together to enable complex functions. We are optimistic that we will make significant progress in this area in the coming decade by undertaking systematic biological and other multidisciplinary approaches.

References

1. Hwang SJ, Park H, Park K. Gastric retentive drug-delivery systems. *Crit Rev Ther Drug Carrier Syst* 1998;**15**:243–84.
2. Jeong EJ, Liu X, Jia X, Chen J, Hu M. Coupling of conjugating enzymes and efflux transporters: impact on bioavailability and drug interactions. *Curr Drug Metab* 2005;**6**:455–68.
3. Hu M, Cheng Z, Zheng L. Functional and molecular characterization of rat intestinal prolidase. *Pediatr Res* 2003;**53**:905–14.
4. Johnson BM, Zhang P, Schuetz JD, Brouwer KL. Characterization of transport protein expression in multidrug resistance-associated protein (Mrp) 2-deficient rats. *Drug Metab Dispos* 2006;**34**:556–62.
5. Klaassen CD, Slitt AL. Regulation of hepatic transporters by xenobiotic receptors. *Curr Drug Metab* 2005;**6**:309–28.
6. Fix JA, Engle K, Porter PA, Leppert PS, Selk SJ, Gardner CR, et al. Acylcarnitines: drug absorption-enhancing agents in the gastrointestinal tract. *Am J Physiol* 1986;**251**:G332–40.
7. Nellans HN. (B) Mechanisms of peptide and protein absorption (1) paracellular intestinal transport: modulation of absorption. *Adv Drug Deliv Rev* 1991;**7**:339–64.
8. Lee VHL, Akira Y. Penetration and enzymatic barriers to peptide and protein absorption. *Adv Drug Deliv Rev* 1989;**4**:171–207.
9. Collington GK, Hunter J, Allen CN, Simmons NL, Hirst BH. Polarized efflux of 2′,7′-bis(2-carboxyethyl)-5(6)-carboxyfluorescein from cultured epithelial cell monolayers. *Biochem Pharmacol* 1992;**44**:417–24.
10. Inui K, Yamamoto M, Saito H. Transepithelial transport of oral cephalosporins by monolayers of intestinal epithelial cell line Caco-2: specific transport systems in apical and basolateral membranes. *J Pharmacol Exp Ther* 1992;**261**:195–201.
11. Hidalgo IJ, Ryan FM, Marks GJ, Smith PL. pH-dependent transepithelial transport of cephalexin in rabbit intestinal mucosa. *Int J Pharm* 1993;**98**:83–92.
12. Caroline H, Gochoco FMR, Miller J, Smith PL, Hidalgo IJ. Uptake and transepithelial transport of the orally absorbed cephalosporin cephalexin, in the human intestinal cell line, Caco-2. *Int J Pharm* 1994;**104**:187–202.
13. Hu M, Amidon GL. Passive and carrier-mediated intestinal absorption components of captopril. *J Pharm Sci* 1988;**77**:1007–11.
14. Dantzig AH, Bergin L. Carrier-mediated uptake of cephalexin in human intestinal cells. *Biochem Biophys Res Commun* 1988;**155**:1082–7.
15. Dantzig AH, Duckworth DC, Tabas LB. Transport mechanisms responsible for the absorption of loracarbef, cefixime, and cefuroxime axetil into human intestinal Caco-2 cells. *Biochim Biophys Acta* 1994;**1191**:7–13.
16. Yang CY, Dantzig AH, Pidgeon C. Intestinal peptide transport systems and oral drug availability. *Pharm Res* 1999;**16**:1331–43.
17. Ganapathy V, Leibach FH. Role of pH gradient and membrane potential in dipeptide transport in intestinal and renal brush-border membrane vesicles from the rabbit. Studies with L-carnosine and glycyl-L-proline. *J Biol Chem* 1983;**258**:14189–92.
18. Ogihara H, Saito H, Shin BC, Terado T, Takenoshita S, Nagamachi Y, et al. Immuno-localization of H^+/peptide cotransporter in rat digestive tract. *Biochem Biophys Res Commun* 1996;**220**:848–52.
19. Rubio-Aliaga I, Boll M, Daniel H. Cloning and characterization of the gene encoding the mouse peptide transporter PEPT2. *Biochem Biophys Res Commun* 2000;**276**:734–41.
20. Fei YJ, Sugawara M, Liu JC, Li HW, Ganapathy V, Ganapathy ME, et al. cDNA structure, genomic organization, and promoter analysis of the mouse intestinal peptide transporter PEPT1. *Biochim Biophys Acta* 2000;**1492**:145–54.
21. Liang R, Fei YJ, Prasad PD, Ramamoorthy S, Han H, Yang-Feng TL, et al. Human intestinal H^+/peptide cotransporter. Cloning, functional expression, and chromosomal localization. *J Biol Chem* 1995;**270**:6456–63.
22. Saito H, Okuda M, Terada T, Sasaki S, Inui K. Cloning and characterization of a rat H^+/peptide cotransporter mediating absorption of beta-lactam antibiotics in the intestine and kidney. *J Pharmacol Exp Ther* 1995;**275**:1631–7.
23. Boll M, Herget M, Wagener M, Weber WM, Markovich D, Biber J, et al. Expression cloning and functional characterization of the kidney cortex high-affinity proton-coupled peptide transporter. *Proc Natl Acad Sci USA* 1996;**93**:284–9.
24. Ramamoorthy S, Liu W, Ma YY, Yang-Feng TL, Ganapathy V, Leibach FH. Proton/peptide cotransporter (PEPT 2) from human kidney: functional characterization and chromosomal localization. *Biochim Biophys Acta* 1995;**1240**:1–4.
25. Terada T, Inui K. Peptide transporters: structure, function, regulation and application for drug delivery. *Curr Drug Metab* 2004;**5**:85–94.
26. Terada T, Sawada K, Ito T, Saito H, Hashimoto Y, Inui K. Functional expression of novel peptide transporter in renal basolateral membranes. *Am J Physiol Renal Physiol* 2000;**279**:F851–7.

27. Sawada K, Terada T, Saito H, Inui K. Distinct transport characteristics of basolateral peptide transporters between MDCK and Caco-2 cells. *Pflugers Arch* 2001;**443**:31–7.
28. Hu M, Subramanian P, Mosberg HI, Amidon GL. Use of the peptide carrier system to improve the intestinal absorption of L-alpha-methyldopa: carrier kinetics, intestinal permeabilities, and in vitro hydrolysis of dipeptidyl derivatives of L-alpha-methyldopa. *Pharm Res* 1989;**6**:66–70.
29. Nakanishi T, Tamai I, Sai Y, Sasaki T, Tsuji A. Carrier-mediated transport of oligopeptides in the human fibrosarcoma cell line HT1080. *Cancer Res* 1997;**57**:4118–22.
30. Mann GE, Yudilevich DL, Sobrevia L. Regulation of amino acid and glucose transporters in endothelial and smooth muscle cells. *Physiol Rev* 2003;**83**:183–252.
31. Qian X, Vinnakota S, Edwards C, Sarkar HK. Molecular characterization of taurine transport in bovine aortic endothelial cells. *Biochim Biophys Acta* 2000;**1509**:324–34.
32. Deves R, Boyd CA. Transporters for cationic amino acids in animal cells: discovery, structure, and function. *Physiol Rev* 1998;**78**:487–545.
33. Cheeseman C. Role of intestinal basolateral membrane in absorption of nutrients. *Am J Physiol* 1992;**263**:R482–8.
34. Cheeseman CI. Characteristics of lysine transport across the serosal pole of the anuran small intestine. *J Physiol* 1983;**338**:87–97.
35. Hammerman MR. Na^+-independent L-arginine transport in rabbit renal brush border membrane vesicles. *Biochim Biophys Acta* 1982;**685**:71–7.
36. Hilden SA, Sacktor B. L-arginine uptake into renal brush border membrane vesicles. *Arch Biochem Biophys* 1981;**210**:289–97.
37. Stieger B, Stange G, Biber J, Murer H. Transport of L-lysine by rat renal brush border membrane vesicles. *Pflugers Arch* 1983;**397**:106–13.
38. You G. The role of organic ion transporters in drug disposition: an update. *Curr Drug Metab* 2004;**5**:55–62.
39. Bulacio RP, Torres AM. Organic anion transporter 5 (Oat5) renal expression and urinary excretion in rats treated with cisplatin: a potential biomarker of cisplatin-induced nephrotoxicity. *Arch Toxicol* 2013;**87**:1953–62.
40. Pritchard JB, Miller DS. Mechanisms mediating renal secretion of organic anions and cations. *Physiol Rev* 1993;**73**:765–96.
41. Sweet DH, Pritchard JB. The molecular biology of renal organic anion and organic cation transporters. *Cell Biochem Biophys* 1999;**31**:89–118.
42. Koepsell H, Gorboulev V, Arndt P. Molecular pharmacology of organic cation transporters in kidney. *J Membr Biol* 1999;**167**:103–17.
43. Tamai I, Yabuuchi H, Nezu J, Sai Y, Oku A, Shimane M, et al. Cloning and characterization of a novel human pH-dependent organic cation transporter, OCTN1. *FEBS Lett* 1997;**419**:107–11.
44. Wu X, George RL, Huang W, Wang H, Conway SJ, Leibach FH, et al. Structural and functional characteristics and tissue distribution pattern of rat OCTN1, an organic cation transporter, cloned from placenta. *Biochim Biophys Acta* 2000;**1466**:315–27.
45. Tamai I, Ohashi R, Nezu JI, Sai Y, Kobayashi D, Oku A, et al. Molecular and functional characterization of organic cation/carnitine transporter family in mice. *J Biol Chem* 2000;**275**:40064–72.
46. Wu X, Huang W, Prasad PD, Seth P, Rajan DP, Leibach FH, et al. Functional characteristics and tissue distribution pattern of organic cation transporter 2 (OCTN2), an organic cation/carnitine transporter. *J Pharmacol Exp Ther* 1999;**290**:1482–92.
47. Wu X, Prasad PD, Leibach FH, Ganapathy V. cDNA sequence, transport function, and genomic organization of human OCTN2, a new member of the organic cation transporter family. *Biochem Biophys Res Commun* 1998;**246**:589–95.
48. Wagner CA, Lukewille U, Kaltenbach S, Moschen I, Broer A, Risler T, et al. Functional and pharmacological characterization of human Na(+)-carnitine cotransporter hOCTN2. *Am J Physiol Renal Physiol* 2000;**279**:F584–91.
49. Cattori V, van Montfoort JE, Stieger B, Landmann L, Meijer DK, Winterhalter KH, et al. Localization of organic anion transporting polypeptide 4 (Oatp4) in rat liver and comparison of its substrate specificity with Oatp1, Oatp2 and Oatp3. *Pflugers Arch* 2001;**443**:188–95.
50. Gao B, Hagenbuch B, Kullak-Ublick GA, Benke D, Aguzzi A, Meier PJ. Organic anion-transporting polypeptides mediate transport of opioid peptides across blood–brain barrier. *J Pharmacol Exp Ther* 2000;**294**:73–9.
51. Kullak-Ublick GA, Ismair MG, Stieger B, Landmann L, Huber R, Pizzagalli F, et al. Organic anion-transporting polypeptide B (OATP-B) and its functional comparison with three other OATPs of human liver. *Gastroenterology* 2001;**120**:525–33.
52. Nozawa T, Tamai I, Sai Y, Nezu J, Tsuji A. Contribution of organic anion transporting polypeptide OATP-C to hepatic elimination of the opioid pentapeptide analogue [D-Ala2, D-Leu5]-enkephalin. *J Pharm Pharmacol* 2003;**55**:1013–20.
53. Lee W, Glaeser H, Smith LH, Roberts RL, Moeckel GW, Gervasini G, et al. Polymorphisms in human organic anion-transporting polypeptide 1A2 (OATP1A2): implications for altered drug disposition and central nervous system drug entry. *J Biol Chem* 2005;**280**:9610–17.
54. Reichel C, Gao B, Van Montfoort J, Cattori V, Rahner C, Hagenbuch B, et al. Localization and function of the organic anion-transporting polypeptide Oatp2 in rat liver. *Gastroenterology* 1999;**117**:688–95.
55. Konig J, Cui Y, Nies AT, Keppler D. Localization and genomic organization of a new hepatocellular organic anion transporting polypeptide. *J Biol Chem* 2000;**275**:23161–8.
56. Konig J, Cui Y, Nies AT, Keppler D. A novel human organic anion transporting polypeptide localized to the basolateral hepatocyte membrane. *Am J Physiol Gastrointest Liver Physiol* 2000;**278**:G156–64.
57. Kong W, Engel K, Wang J. Mammalian nucleoside transporters. *Curr Drug Metab* 2004;**5**:63–84.
58. Griffith DA, Jarvis SM. Nucleoside and nucleobase transport systems of mammalian cells. *Biochim Biophys Acta* 1996;**1286**:153–81.
59. Cabrita MA, Baldwin SA, Young JD, Cass CE. Molecular biology and regulation of nucleoside and nucleobase transporter proteins in eukaryotes and prokaryotes. *Biochem Cell Biol* 2002;**80**:623–38.
60. Sue Masters B, Marohnic CC. Cytochromes P450--a family of proteins and scientists-understanding their relationships. *Drug Metab Rev* 2006;**38**:209–25.
61. Jakoby WB. *Detoxication and drug metabolism: conjugation and related systems*. New York, NY: Academic Press; 1981.
62. Testa B, Jenner P. *Drug metabolism: chemical and biochemical aspects*. New York, NY: M. Dekker; 1976.
63. Numa S. *Fatty acid metabolism and its regulation*. Amsterdam; New York, NY: Elsevier; 1984.
64. Conley A, Hinshelwood M. Mammalian aromatases. *Reproduction* 2001;**121**:685–95.
65. Fisher MB, Campanale K, Ackermann BL, VandenBranden M, Wrighton SA. In vitro glucuronidation using human liver microsomes and the pore-forming peptide alamethicin. *Drug Metab Dispos* 2000;**28**:560–6.
66. Guengerich FP. Cytochrome P450s and other enzymes in drug metabolism and toxicity. *AAPS J* 2006;**8**:E101–11.

67. Dervieux T, Meshkin B, Neri B. Pharmacogenetic testing: proofs of principle and pharmacoeconomic implications. *Mutat Res* 2005;**573**:180–94.
68. Bischoff SC. Food allergies. *Curr Gastroenterol Rep* 2006;**8**:374–82.
69. Liu Y, Hu M. Absorption and metabolism of flavonoids in the Caco-2 cell culture model and a perused rat intestinal model. *Drug Metab Dispos* 2002;**30**:370–7.
70. Liu Y, Liu Y, Dai Y, Xun L, Hu M. Enteric disposition and recycling of flavonoids and ginkgo flavonoids. *J Altern Complement Med* 2003;**9**:631–40.
71. Chen J, Lin H, Hu M. Metabolism of flavonoids via enteric recycling: role of intestinal disposition. *J Pharmacol Exp Ther* 2003;**304**:1228–35.
72. Hu M, Krausz K, Chen J, Ge X, Li J, Gelboin HL, et al. Identification of CYP1A2 as the main isoform for the phase I hydroxylated metabolism of genistein and a prodrug converting enzyme of methylated isoflavones. *Drug Metab Dispos* 2003;**31**:924–31.
73. Wang SW, Chen J, Jia X, Tam VH, Hu M. Disposition of flavonoids via enteric recycling: structural effects and lack of correlations between in vitro and in situ metabolic properties. *Drug Metab Dispos* 2006;**34**:1837–48.
74. Rowland A, Miners JO, Mackenzie PI. The UDP-glucuronosyltransferases: their role in drug metabolism and detoxification. *Int J Biochem Cell Biol* 2013;**45**:1121–32.
75. Mackenzie PI, Bock KW, Burchell B, Guillemette C, Ikushiro S, Iyanagi T, et al. Nomenclature update for the mammalian UDP glycosyltransferase (UGT) gene superfamily. *Pharmacogenet Genomics* 2005;**15**:677–85.
76. Gamage N, Barnett A, Hempel N, Duggleby RG, Windmill KF, Martin JL, et al. Human sulfotransferases and their role in chemical metabolism. *Toxicol Sci* 2006;**90**:5–22.
77. Nimmagadda D, Cherala G, Ghatta S. Cytosolic sulfotransferases. *Indian J Exp Biol* 2006;**44**:171–82.
78. Walle T, Hsieh F, DeLegge MH, Oatis Jr JE, Walle UK. High absorption but very low bioavailability of oral resveratrol in humans. *Drug Metab Dispos* 2004;**32**:1377–82.
79. Hayes JD, Flanagan JU, Jowsey IR. Glutathione transferases. *Annu Rev Pharmacol Toxicol* 2005;**45**:51–88.
80. Haga SB, Thummel KE, Burke W. Adding pharmacogenetics information to drug labels: lessons learned. *Pharmacogenet Genomics* 2006;**16**:847–54.
81. Coles BF, Kadlubar FF. Detoxification of electrophilic compounds by glutathione S-transferase catalysis: determinants of individual response to chemical carcinogens and chemotherapeutic drugs? *Biofactors* 2003;**17**:115–30.
82. Kessel D, Botterill V, Wodinsky I. Uptake and retention of daunomycin by mouse leukemic cells as factors in drug response. *Cancer Res* 1968;**28**:938–41.
83. Gottesman MM, Ling V. The molecular basis of multidrug resistance in cancer: the early years of P-glycoprotein research. *FEBS Lett* 2006;**580**:998–1009.
84. Hyde SC, Emsley P, Hartshorn MJ, Mimmack MM, Gileadi U, Pearce SR, et al. Structural model of ATP-binding proteins associated with cystic fibrosis, multidrug resistance and bacterial transport. *Nature* 1990;**346**:362–5.
85. Evans AM. Influence of dietary components on the gastrointestinal metabolism and transport of drugs. *Ther Drug Monit* 2000;**22**:131–6.
86. Hunter J, Hirst BH, Simmons NL. Epithelial secretion of vinblastine by human intestinal adenocarcinoma cell (HCT-8 and T84) layers expressing P-glycoprotein. *Br J Cancer* 1991;**64**:437–44.
87. Schinkel AH, Mayer U, Wagenaar E, Mol CA, van Deemter L, Smit JJ, et al. Normal viability and altered pharmacokinetics in mice lacking mdr1-type (drug-transporting) P-glycoproteins. *Proc Natl Acad Sci USA* 1997;**94**:4028–33.
88. Cole SP, Bhardwaj G, Gerlach JH, Mackie JE, Grant CE, Almquist KC, et al. Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line. *Science* 1992;**258**:1650–4.
89. Buchler M, Konig J, Brom M, Kartenbeck J, Spring H, Horie T, et al. cDNA cloning of the hepatocyte canalicular isoform of the multidrug resistance protein, cMrp, reveals a novel conjugate export pump deficient in hyperbilirubinemic mutant rats. *J Biol Chem* 1996;**271**:15091–8.
90. Paulusma CC, Bosma PJ, Zaman GJ, Bakker CT, Otter M, Scheffer GL, et al. Congenital jaundice in rats with a mutation in a multidrug resistance-associated protein gene. *Science* 1996;**271**:1126–8.
91. Belinsky MG, Kruh GD. MOAT-E (ARA) is a full-length MRP/cMOAT subfamily transporter expressed in kidney and liver. *Br J Cancer* 1999;**80**:1342–9.
92. Stride BD, Grant CE, Loe DW, Hipfner DR, Cole SP, Deeley RG. Pharmacological characterization of the murine and human orthologs of multidrug-resistance protein in transfected human embryonic kidney cells. *Mol Pharmacol* 1997;**52**:344–53.
93. de Jong MC, Slootstra JW, Scheffer GL, Schroeijers AB, Puijk WC, Dinkelberg R, et al. Peptide transport by the multidrug resistance protein MRP1. *Cancer Res* 2001;**61**:2552–7.
94. Cole SP, Sparks KE, Fraser K, Loe DW, Grant CE, Wilson GM, et al. Pharmacological characterization of multidrug resistant MRP-transfected human tumor cells. *Cancer Res* 1994;**54**:5902–10.
95. Kawabe T, Chen ZS, Wada M, Uchiumi T, Ono M, Akiyama S, et al. Enhanced transport of anticancer agents and leukotriene C4 by the human canalicular multispecific organic anion transporter (cMOAT/MRP2). *FEBS Lett* 1999;**456**:327–31.
96. Paulusma CC, Oude Elferink RP. The canalicular multispecific organic anion transporter and conjugated hyperbilirubinemia in rat and man. *J Mol Med* 1997;**75**:420–8.
97. Chen Y, Zhao YH, Jia XB, Hu M. Intestinal absorption mechanisms of prenylated flavonoids present in the heat-processed Epimedium koreanum Nakai (Yin Yanghuo). *Pharm Res* 2008;**25**:2190–9.
98. Xu H, Kulkarni KH, Singh R, Yang Z, Wang SW, Tam VH, et al. Disposition of naringenin via glucuronidation pathway is affected by compensating efflux transporters of hydrophilic glucuronides. *Mol Pharm* 2009;**6**:1703–15.
99. Liu W, Feng Q, Li Y, Ye L, Hu M, Liu Z. Coupling of UDP-glucuronosyltransferases and multidrug resistance-associated proteins is responsible for the intestinal disposition and poor bioavailability of emodin. *Toxicol Appl Pharmacol* 2012;**265**:316–24.
100. Ge S, Yin T, Xu B, Gao S, Hu M. Curcumin affects phase II disposition of resveratrol through inhibiting efflux transporters MRP2 and BCRP. *Pharm Res* 2015.
101. van Herwaarden AE, Schinkel AH. The function of breast cancer resistance protein in epithelial barriers, stem cells and milk secretion of drugs and xenotoxins. *Trends Pharmacol Sci* 2006;**27**:10–16.
102. Vlaming ML, Lagas JS, Schinkel AH. Physiological and pharmacological roles of ABCG2 (BCRP): recent findings in Abcg2 knockout mice. *Adv Drug Deliv Rev* 2009;**61**:14–25.
103. Wakabayashi K, Tamura A, Saito H, Onishi Y, Ishikawa T. Human ABC transporter ABCG2 in xenobiotic protection and redox biology. *Drug Metab Rev* 2006;**38**:371–91.
104. Mao Q, Unadkat JD. Role of the breast cancer resistance protein (BCRP/ABCG2) in drug transport--an update. *AAPS J* 2015;**17**:65–82.
105. Ni Z, Bikadi Z, Rosenberg MF, Mao Q. Structure and function of the human breast cancer resistance protein (BCRP/ABCG2). *Curr Drug Metab* 2010;**11**:603–17.

106. Allikmets R, Schriml LM, Hutchinson A, Romano-Spica V, Dean M. A human placenta-specific ATP-binding cassette gene (ABCP) on chromosome 4q22 that is involved in multidrug resistance. *Cancer Res* 1998;**58**:5337–9.
107. Miyake K, Mickley L, Litman T, Zhan Z, Robey R, Cristensen B, et al. Molecular cloning of cDNAs which are highly overexpressed in mitoxantrone-resistant cells: demonstration of homology to ABC transport genes. *Cancer Res* 1999;**59**:8–13.
108. Stacy AE, Jansson PJ, Richardson DR. Molecular pharmacology of ABCG2 and its role in chemoresistance. *Mol Pharmacol* 2013;**84**:655–69.
109. Robey RW, Steadman K, Polgar O, Bates SE. ABCG2-mediated transport of photosensitizers: potential impact on photodynamic therapy. *Cancer Biol Ther* 2005;**4**:187–94.
110. Sekine T, Watanabe N, Hosoyamada M, Kanai Y, Endou H. Expression cloning and characterization of a novel multispecific organic anion transporter. *J Biol Chem* 1997;**272**:18526–9.
111. Sweet DH, Wolff NA, Pritchard JB. Expression cloning and characterization of ROAT1. The basolateral organic anion transporter in rat kidney. *J Biol Chem* 1997;**272**:30088–95.
112. Moller JV, Sheikh MI. Renal organic anion transport system: pharmacological, physiological, and biochemical aspects. *Pharmacol Rev* 1982;**34**:315–58.
113. Sweet DH, Bush KT, Nigam SK. The organic anion transporter family: from physiology to ontogeny and the clinic. *Am J Physiol Renal Physiol* 2001;**281**:F197–205.
114. Deguchi T, Kusuhara H, Takadate A, Endou H, Otagiri M, Sugiyama Y. Characterization of uremic toxin transport by organic anion transporters in the kidney. *Kidney Int* 2004;**65**:162–74.
115. Benet LZ, Cummins CL, Wu CY. Unmasking the dynamic interplay between efflux transporters and metabolic enzymes. *Int J Pharm* 2004;**277**:3–9.
116. Cummins CL, Jacobsen W, Benet LZ. Unmasking the dynamic interplay between intestinal P-glycoprotein and CYP3A4. *J Pharmacol Exp Ther* 2002;**300**:1036–45.
117. Coldham NG, Zhang AQ, Key P, Sauer MJ. Absolute bioavailability of [14C] genistein in the rat; plasma pharmacokinetics of parent compound, genistein glucuronide and total radioactivity. *Eur J Drug Metab Pharmacokinet* 2002;**27**:249–58.
118. Setchell KD, Brown NM, Desai PB, Zimmer-Nechimias L, Wolfe B, Jakate AS, et al. Bioavailability, disposition, and dose-response effects of soy isoflavones when consumed by healthy women at physiologically typical dietary intakes. *J Nutr* 2003;**133**:1027–35.
119. Wang SW, Chen J, Jia X, Tam VH, Hu M. Disposition of flavonoids via enteric recycling: structural effects and lack of correlations between in vitro and in situ metabolic properties. *Drug Metab Dispos* 2006.
120. Chen J, Lin H, Hu M. Metabolism of flavonoids via enteric recycling: role of intestinal disposition. *J Pharmacol Exp Ther* 2003;**304**:1228–35.
121. Ito K, Suzuki H, Horie T, Sugiyama Y. Apical/basolateral surface expression of drug transporters and its role in vectorial drug transport. *Pharm Res* 2005;**22**:1559–77.
122. Ma Y, Zeng M, Sun R, Hu M. Disposition of flavonoids impacts their efficacy and safety. *Curr Drug Metab* 2014;**15**:841–64.
123. Matsukawa N, Matsumoto M, Hara H. High biliary excretion levels of quercetin metabolites after administration of a quercetin glycoside in conscious bile duct cannulated rats. *Biosci Biotechnol Biochem* 2009;**73**:1863–5.
124. Jia X, Chen J, Lin H, Hu M. Disposition of flavonoids via enteric recycling: enzyme-transporter coupling affects metabolism of biochanin A and formononetin and excretion of their phase II conjugates. *J Pharmacol Exp Ther* 2004;**310**:1103–13.
125. Prasain JK, Xu J, Kirk M, Johnson MS, Sfakianos J, Barnes S. Differential biliary excretion of genistein metabolites following intraduodenal and intravenous infusion of genistin in female rats. *J Nutr* 2006;**136**:2975–9.
126. Walle T. Absorption and metabolism of flavonoids. *Free Radic Biol Med* 2004;**36**:829–37.
127. McKenna NJ, Lanz RB, O'Malley BW. Nuclear receptor coregulators: cellular and molecular biology. *Endocr Rev* 1999;**20**:321–44.
128. Nwachukwu JC, Nettles KW. The nuclear receptor signalling scaffold: insights from full-length structures. *EMBO J* 2012;**31**:251–3.
129. Chai X, Zeng S, Xie W. Nuclear receptors PXR and CAR: implications for drug metabolism regulation, pharmacogenomics and beyond. *Expert Opin Drug Metab Toxicol* 2013;**9**:253–66.
130. Teotico DG, Frazier ML, Ding F, Dokholyan NV, Temple BR, Redinbo MR. Active nuclear receptors exhibit highly correlated AF-2 domain motions. *PLoS Comput Biol* 2008;**4**:e1000111.
131. Goodwin B, Hodgson E, Liddle C. The orphan human pregnane X receptor mediates the transcriptional activation of CYP3A4 by rifampicin through a distal enhancer module. *Mol Pharmacol* 1999;**56**:1329–39.
132. Lehmann JM, McKee DD, Watson MA, Willson TM, Moore JT, Kliewer SA. The human orphan nuclear receptor PXR is activated by compounds that regulate CYP3A4 gene expression and cause drug interactions. *J Clin Invest*. 1998;**102**:1016–23.
133. Bertilsson G, Heidrich J, Svensson K, Asman M, Jendeberg L, Sydow-Backman M, et al. Identification of a human nuclear receptor defines a new signaling pathway for CYP3A induction. *Proc Natl Acad Sci USA* 1998;**95**:12208–13.
134. Ingraham HA, Redinbo MR. Orphan nuclear receptors adopted by crystallography. *Curr Opin Struct Biol* 2005;**15**:708–15.
135. Sugatani J, Kojima H, Ueda A, Kakizaki S, Yoshinari K, Gong QH, et al. The phenobarbital response enhancer module in the human bilirubin UDP-glucuronosyltransferase UGT1A1 gene and regulation by the nuclear receptor CAR. *Hepatology (Baltimore, MD)* 2001;**33**:1232–8.
136. Nims RW, Sinclair PR, Sinclair JF, Dragnev KH, Jones CR, Mellini DW, et al. Dose-response relationships for the induction of P450 2B by 1,4-bis[2-(3,5-dichloropyridyloxy)]benzne (TCPOBOP) in rat and cultured rat hepatocytes. *Xenobiotica* 1993;**23**:1411–26.
137. Lu L, Shi J, Li Q, Peng X, Dong L, Li Y, et al. Effects of liver diseases on drug-metabolizing enzymes: implications for drug fate alterations and nano-therapeutic openings. *Curr Med Chem* 2014;**21**:2522–41.
138. Hanada K, Nakai K, Tanaka H, Suzuki F, Kumada H, Ohno Y, et al. Effect of nuclear receptor downregulation on hepatic expression of cytochrome P450 and transporters in chronic hepatitis C in association with fibrosis development. *Drug Metab Pharmacokinet* 2012;**27**:301–6.
139. Senekeo-Effenberger K, Chen S, Brace-Sinnokrak E, Bonzo JA, Yueh MF, Argikar U, et al. Expression of the human UGT1 locus in transgenic mice by 4-chloro-6-(2,3-xylidino)-2-pyrimidinylthioacetic acid (WY-14643) and implications on drug metabolism through peroxisome proliferator-activated receptor alpha activation. *Drug Metab Dispos* 2007;**35**:419–27.
140. Tachibana S, Yoshinari K, Chikada T, Toriyabe T, Nagata K, Yamazoe Y. Involvement of Vitamin D receptor in the intestinal induction of human ABCB1. *Drug Metab Dispos* 2009;**37**:1604–10.
141. Schmiedlin-Ren P, Thummel KE, Fisher JM, Paine MF, Watkins PB. Induction of CYP3A4 by 1 alpha,25-dihydroxyvitamin D3 is human cell line-specific and is unlikely to involve pregnane X receptor. *Drug Metab Dispos* 2001;**29**:1446–53.

142. Jeong HS, Ryoo IG, Kwak MK. Regulation of the expression of renal drug transporters in KEAP1-knockdown human tubular cells. *Toxicol In Vitro* 2015;**29**:884–92.
143. Hoque MT, Robillard KR, Bendayan R. Regulation of breast cancer resistant protein by peroxisome proliferator-activated receptor alpha in human brain microvessel endothelial cells. *Mol Pharmacol* 2012;**81**:598–609.
144. Lipinski CA. Chris Lipinski discusses life and chemistry after the Rule of Five. *Drug Discov Today* 2003;**8**:12–16.
145. Remko M, Swart M, Bickelhaupt FM. Theoretical study of structure, pK_a, lipophilicity, solubility, absorption, and polar surface area of some centrally acting antihypertensives. *Bioorg Med Chem* 2006;**14**:1715–28.
146. Wenlock MC, Austin RP, Barton P, Davis AM, Leeson PD. A comparison of physiochemical property profiles of development and marketed oral drugs. *J Med Chem* 2003;**46**:1250–6.
147. Dunn CJ, Wagstaff AJ, Perry CM, Plosker GL, Goa KL. Cyclosporin: an updated review of the pharmacokinetic properties, clinical efficacy and tolerability of a microemulsion-based formulation (neoral)1 in organ transplantation. *Drugs* 2001;**61**:1957–2016.
148. Constable PD. Acid-base assessment: when and how to apply the Henderson-Hasselbalch equation and strong ion difference theory. *Vet Clin North Am Food Anim Pract* 2014;**30**:295–316 v.
149. Song N-N, Zhang S-Y, Liu C-X. Overview of factors affecting oral drug absorption. *Asian J Drug Metab Pharmacokinet* 2004;**4**:167–76.
150. Aulton ME, Cooper JW. *Pharmaceutics: the science of dosage form design*. 1st ed. Edinburgh; New York, NY: Churchill Livingstone; 1988.
151. Kimura T, Higaki K. Gastrointestinal transit and drug absorption. *Biol Pharm Bull* 2002;**25**:149–64.
152. Shargel L, Wu-Pong S, Yu ABC. Medical Pub. Division *Applied biopharmaceutics & pharmacokinetics*. 5th ed. New York, NY: McGraw-Hill; 2005.
153. Martinez MN, Amidon GL. A mechanistic approach to understanding the factors affecting drug absorption: a review of fundamentals. *J Clin Pharmacol* 2002;**42**:620–43.
154. Ritschel WA, Kearns GL, American Pharmaceutical Association. *Handbook of basic pharmacokinetics--including clinical applications*. 6th ed. Washington, DC: American Pharmacists Association; 2004.

This page intentionally left blank

CHAPTER

12

Oral Drug Absorption: Evaluation and Prediction

Y. Yang[1], *Y. Zhao*[1], *A. Yu*[2], *D. Sun*[2] *and L.X. Yu*[1]

[1]U.S. Food and Drug Administration, Silver Spring, MD, United States
[2]University of Michigan, Ann Arbor, MI, United States

12.1 INTRODUCTION

The most convenient route of drug administration for systemic actions is oral. The most important question during drug development for an oral-dosage form is how much of a drug will be bioavailable after its oral administration, and how quickly that happens. Successful design and development of orally bioavailable drugs requires that the drug molecules (1) be reasonably stable to chemical and enzymatic degradation (eg, within the intestinal lumen, intestinal wall, liver, and circulation); (2) be able to traverse the intestinal epithelial barrier to the portal circulation; and (3) can pass through the liver and enter the systemic circulation intact at a sufficient concentration to elicit the appropriate pharmacologic response. Therefore, good oral bioavailability generally facilitates formulation design and development, reduces intrasubject variability, and enhances dosing flexibility. Oral bioavailability (F) can be mathematically represented by

$$F = f_a \times f_g \times f_h \tag{12.1}$$

where, f_a is the fraction of drug absorbed, f_g is the fraction that escapes metabolism in the gastrointestinal (GI) tract, and f_h is the fraction that escapes the first-pass hepatic metabolism. Each individual value of f_a, f_g, or f_h ranges from 0 (completely unabsorbed or metabolized) to 1 (fully absorbed or unmetabolized). Therefore, oral drug absorption is one of the main factors governing drug bioavailability.

Oral drug absorption is the movement of the drug from its site of administration, the GI tract, into the bloodstream. The oral absorption of the drug in the solid dosage form from the GI tract is largely controlled by (1) dissolution rate and solubility, which determine how fast a drug reaches maximum concentration in the GI fluid; and (2) intestinal permeability, which relates to the rate at which a dissolved drug crosses the intestinal wall to reach the portal blood circulation. The determination of the dissolution, solubility, and permeability properties of drugs can thus provide information about drug absorption. This chapter will focus on these properties. It begins with the GI physiology. It then discusses the biopharmaceutical classification system (BCS). Further, it proceeds to introduce in vitro, in silico, and in vivo methods to evaluate intestinal permeability or to predict oral drug absorption. The chapter concludes by discussing future trends in oral drug absorption evaluation and prediction.

12.2 ANATOMY AND PHYSIOLOGY OF THE GI TRACT

The GI tract begins at the mouth, includes the esophagus, stomach, small intestine, large intestine (also known as the colon), and rectum, and ends at the anus. The entire system is about 9 m long in the average human being, according to the American Society of Gastrointestinal Endoscopy[1] and the entire GI tract is lined with mucous membrane, through which drugs may be readily transferred into the systemic circulation, and is highly perfused by a capillary network, which makes drug absorption possible. The human GI system absorbs nutrients and fluids in parallel with its

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00012-1

function as an efficient barrier against potentially hazardous bacteria and ingested and locally formed toxins. After a drug is given orally, GI transit and environments with respect to the surface area and membrane, pH, property of the GI fluids, enzymes, or intestinal transporters interact with drug molecules and influence drug absorption.

Small intestine (duodenum, jejunum, and upper region of ileum), the major and most efficient site for drug absorption of most drugs, has certain specific features: (1) an overall large surface area (0.33 m^2); (2) folds that result in a threefold increase in surface area (1 m^2); (3) fingerlike projections called *villi*, which increase the surface area by 30 times (10 m^2); (4) microvilli, which increase the surface area by 600 times (200 m^2)[2]; (5) a pH range of 5–7, which is favorable for most drugs to remain unionized; (6) a rich capillary network; and (7) transit time (3–4 hours).

The large intestine has specific features: (1) the presence of microbial flora localized in the colon, and the amount of bacteria varies depending on diet and use of antibiotics but can make up more than half the weight of fecal solids[3]; (2) a near-neutral pH; (3) reduced digestive enzymatic activity; (4) a much greater responsiveness to absorption enhancers; and (5) a much longer transit time (35–65 hours).[4]

Gastric residence time is variable, depending on many physiological, pathological, and pharmaceutical factors. For example, body posture, age, gender, osmolality, pH, and food are capable of influencing the gastric emptying. Fasted state gastric motility is influenced by the Interdigestive migrating motor complex (MMC).[5] There are four phases associated with the MMC. Phase I is approximately 45–60 minutes long and showcases minimal activity with the slowest corresponding gastric emptying. Phase II is 15–45 minutes long and consists of intermittent and erratic contractions. Phase III is the shortest phase, consisting of intense, short bursts of behavior lasting 5–10 minutes. Phase IV is a transition period from phase III back to phase I, lasting approximately 15–30 minutes.

The reported pH values of fasted gastric, duodenal, and jejunal fluids could range from 1.7 to 3.3 (median of 2.5) in stomach, 5.6–7.0 (median of 6.3) in duodenum, and 6.5–7.8 (median of 6.9) in jejunum.[6] The GI pH under the fed state will largely be governed by the administered meal itself, and the pH of the intestinal fluids is not affected to the same extent as gastric fluids because of the pretreatment in the stomach.[6] Under fed conditions, the duodenal pH is reported to be 5.4–6.5 (median of 6.0),[6] which is overlapping the fasted duodenal pH. Under fed conditions, the jejunal pH is reported to be 6.1.[6]

About 8–10 L of GI fluids are secreted into the GI tract daily. Gastric fluid is a complex and dynamic mixture consisting of saliva, gastric secretions, food, liquid intake, and refluxed liquid from duodenum. A literature review of parameters that characterize some attributes of gastric fluid and its subsequent intestinal presence is detailed in Tables 12.1 and 12.2 (reproduced from Mudie et al.[7]). Magnetic resonance imaging clinical studies have observed the formation and existence of small, discrete fluid pockets in the small intestine.[8] This fluid is also not distributed homogenously (as shown in Figs. 12.1 and 12.2). This method has been utilized to quantify GI fluid volume and transit after ingestion of 240 mL of water in fasted volunteers (as shown in Fig. 12.3).[9]

12.3 BIOPHARMACEUTICS CLASSIFICATION SYSTEM

The BCS is a scientific framework for classifying a drug substance based on its aqueous solubility and intestinal permeability.[10] According to this system, drug substances can be classified into four groups: (I) high solubility–high permeability, (II) low solubility–high permeability, (III) high solubility–low permeability, and (IV) low solubility–low permeability (Table 12.3). When combined with the in vitro dissolution characteristics of the drug product, the BCS considers three major factors: solubility, dissolution rate, and intestinal permeability, all of which govern the rate and extent of oral absorption from immediate release (IR) solid oral-dosage forms. Since the introduction of the BCS, its validity and applicability have been the subject of extensive research and discussion. The BCS framework has implications in the selection of candidate drugs for full development, prediction, and elucidation of food interaction, choice of formulation principle, including suitability for oral extended-release administration,[11] and the possibility of in vitro/in vivo correlation in the dissolution testing of solid formulations.[12] The regulatory applications of BCS are demonstrated by the U.S. Food and Drug Administration (FDA)[13,14] and a 2015 updated draft guidance for industry on waiver of in vivo bioavailability (BA) and bioequivalence (BE) studies for BCS class I/class III drugs in rapid dissolution IR solid oral-dosage forms.[15]

12.3.1 FDA guidance on biowaivers

The 2015 updated guidance[15] provides recommendations for sponsors of investigational new drug (IND) applications, new drug applications (NDAs), abbreviated new drug applications (ANDAs), and supplements to these applications that may request biowaivers

TABLE 12.1 Literature Values for Concentrations of Some Major Components of Fluid in the Fasted and Fed Stomach and Small Intestine

Fluid component			Stomach	Duodenum	Jejunum	Ileum
Bicarbonate (mEg/L^{-1})	Fasted	Mean	7.3	2.7, 6.7, 15	17, 30, 30, 8.2 ± 5 mM	40, 50, 70, 74, 75, 30 ± 11 mM
		Range	9–20		2–20, 5–10, 6–20	
	Fed	Mean		10	2.7, 2.6	
Bile Salts (mM)	Fasted	Median	0.1			
		Mean	0.08 ± 0.03, 0.3, 0.1	6.4 ± 1.3, 4.3 ± 1.2, 5.9 ± 1.8	2 ± 0.2	
		Range		1–5.3, 0.6–5.1, 0.3–9.6	0.8–5.5, 0.1–13.3, 5–6, 0–17	2–10
	Fed	Median		3.6, 5.2, 8.3, 11.9, 11.2, 5.2		1, 0.5
		Mean	0.06	14.5, 5.2, 16 ± 2, 10 ± 1, 9	8, 15, 8 ± 0.1, 6.5 ± 0.9	
		Range		1.6–6.2, 3.2–6.8, 6.7–13.4	0.5–40, 3–34	0.5–30, 0.2–1.3
Lipids (mg/mL)	Fasted	Median		0.5		
		Mean	0.56	0.6	0.1 ± 0.01 mM	
		Range		0–1.8		
	Fed	Median		1.8, 2.6		
		Mean			22 ± 1 mM	
		Range	50, 150	0.5–4.6, 1.1–3.6, 55–100		
Phospholipids (mM)	Fasted	Median		0.6		
		Mean			0.2 ± 0.07	
		Range		0.1–1.5, 0.03–0.06		
	Fed	Median		1.8, 1.2		
		Mean			3 ± 0.3	
		Range		1.3–2.4, 0.8–1.6		
Pepsin	Fasted	Median	0.1, 0.2			
		Mean	0.9			
		Range	0.8–1.3			
	Fed	Median				
		Mean	1.3, 1.7			
		Range	0.3–0.6, 0.6–1.7			
Lipase	Fasted	Mean	About 0.1 mg/mL			
	Fed	Range	11.4–43.9 U/mL			
Potassium (mM)	Fasted	Mean	13.4 ± 3		5.4 ± 2.1, 4.8 ± 0.5	4.9 ± 1.5
Sodium (mM)	Fasted	Mean	68 ± 29		142 ± 13, 142 ± 7	140 ± 6
	Fed	Mean			106 ± 15, 101 ± 17	139 ± 11, 133 ± 8
Chloride (mM)	Fasted	Mean	102 ± 28		126 ± 19, 135 ± 8	125 ± 12
Calcium (mM)	Fasted	Mean	0.6 ± 0.2		0.5 ± 0.3	

Unless indicated next to the value, units are noted next to the name of the component.
This table modified with the permission of American Chemical Society (Mudie DM, Amidon GL, Amidon GE. Physiological parameters for oral delivery and in vitro testing. Mol Pharm *2010;7:1388–405, http://pubs.acs.org/doi/abs/10.1021/mp100149j).*

TABLE 12.2 Literature Values for Properties of Fluids in the Fasted and Fed Stomach and Small Intestine

Fluid property			Stomach	Duodenum	Jejunum	Ileum
Buffer capacity (mmol/L per pH)	Fasted	Median	7, 18	5.6		
		Mean			3.2	6.4
		Range		4–13	2.4–2.8	
	Fed	Range	14–28	18–30	13.2–14.6	
Osmolality (mOsm/kg)	Fasted	Median	98, 140	178, 224		
		Mean	29, 191 ± 36, 33.6 ± 5.9, 221 ± 15	142, 137 ± 54	271 ± 15, 200 ± 68, 278 ± 16	
		Range	171–276	124–266		
	Fed	Median	559, 217	287, 276, > 287, 287		
		Range		250–367, 268–304		
Surface tension (mN/m)	Fasted	Median		32.3, 41.2		
		Mean			28 ± 1, 33.7 ± 2.8	
		Range	41.9–45.7	33.3–46.0		
	Fed	Median		34.2, 35.4		
		Mean			27 ± 1	
		Range	30–31	32.2–36.7, 33.7–36.0		
Viscosity (cP)	Fed	Range	10–2000			
pH	Fasted	Median	1.7, 2.4, 1.7, 1.8	6.1, 6.2, 6.6, 5.6	7.2	
		Mean	2.9 ± 2.0	6.7 ± 0.4, 7.0 ± 0.4, 4.9, 6.4 ± 0.6	6.8 ± 0.4, 7.5, 7.1 ± 0.60	6.5 ± 0.2
		Range	1–2.5, 1.4–2.1, 1.2–7.4, 1.4–7.5	5.8–6.5, 4.0–5.4, 5.2–6.1	4.4–6.5, 5.3–8.1, 5.3–8.1	6.8–8
	Fed	Median	5, 6.4, 2.7	5.4, 6.6, 5.2, 5.9, 6.1, 5.4		
		Mean		5.2, 4.2	6.2 ± 0.2, 5.4 ± 0.2, 6.1	7.5
		Range	4.3–5.4	3.1–6.7, 4.5–5.5, 3.9–4.8, 5.1–5.7, 5.3–6.1, 4.6–6.3	5.2–6.0	6.8–7.8, 6.8–8.0

This table modified with the permission of American Chemical Society (Mudie DM, Amidon GL, Amidon GE. Physiological parameters for oral delivery and in vitro testing. Mol Pharm *2010;7:1388–405, http://pubs.acs.org/doi/abs/10.1021/mp100149j).*

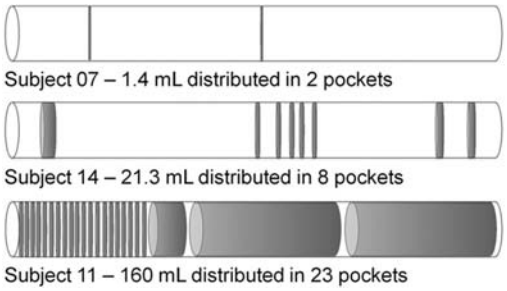


FIGURE 12.1 Graphical representation of the number and volume of small bowel water pockets for three subjects. The small intestine is represented as a cylinder. Shaded areas represent liquid pockets. The minimum (1.4 mL, Subject 14), near-median (21.3 mL, Subject 07), and maximum (160.1 mL, Subject 11) small bowel water volumes at the fasting baseline (before administration of the study dose of water) are shown to illustrate schematically high interindividual variability. *This figure reproduced with the permission of American Chemical Society (Mudie DM, Murray K, Hoad CL, Pritchard SE, Garnett MC, Amidon GL, et al. Quantification of gastrointestinal liquid volumes and distribution following a 240 mL dose of water in the fasted state.* Mol Pharm *2014;**11**:3039–47, http://pubs.acs.org/doi/abs/10.1021/mp500210c).*

for in vivo bioavailability/bioequivalence studies for highly soluble and highly permeable drug substances (class I) with rapid in vitro dissolution, or highly soluble and low-permeable drug substances (class III) with very rapid in vitro dissolution in IR solid oral-dosage forms, provided that the following conditions are met: (1) the drug must be stable in the GI; (2) excipients used in the IR solid oral-dosage forms have no significant effect on the rate and extent of oral drug absorption; (3) the drug must not have a narrow therapeutic index; and (4) the product is designed not to be absorbed in the oral cavity. The guidance also outlines methods for the determination of solubility, dissolution, and permeability.

12.3.1.1 Determination of drug solubility

According to the updated 2015 FDA draft guidance,[15] the pH-solubility profile of the test drug substance should be established at 37 ± 1°C in aqueous media with a pH in the range of 1–6.8; the pH conditions for drug solubility determination should be determined at pH = pK_a, pH = pK_a + 1, pH = pK_a–1, and at pH = 1 and 6.8, with a minimum of three replicate determinations of solubility in each pH condition. Standard buffer solutions used for solubility studies are described in the United States Pharmacopeia and the National Formulary (USP–NF). If these buffers are not suitable for physical or chemical reasons, other buffer solutions can be used. It is necessary to verify the solution pH after addition of the drug substance to a buffer. Methods include the traditional shake-flask method and acid or base titration methods. A validated

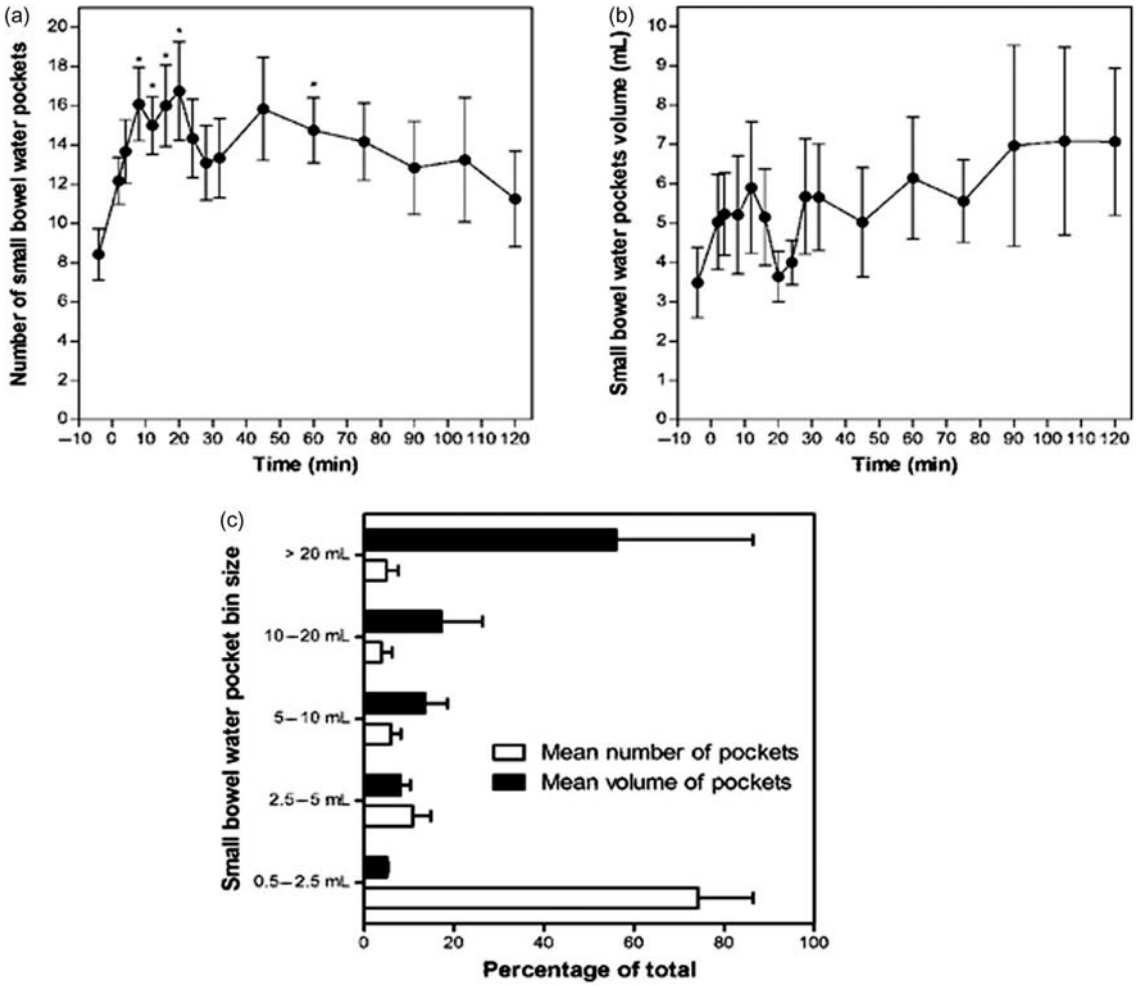


FIGURE 12.2 Small bowel water pockets before and after ingestion of a 240-mL water dose administered at $t = 0$ min (excluding all pockets in the smallest range < 0.5 mL). (a) Mean total number of small bowel water pockets. (b) Mean volume of small bowel water pockets. (c) Comparison between the percentage of the total number of pockets and the percentage of the total volume of each pocket in each volume range at the fasting baseline (before administration of the study drink of water). $n = 12$ healthy volunteers. Error bars represent ±SEM. Dunn's multiple comparison test versus baseline value $^*P < 0.05$. *This figure reproduced with the permission of American Chemical Society (Mudie DM, Murray K, Hoad CL, Pritchard SE, Garnett MC, Amidon GL, et al. Quantification of gastrointestinal liquid volumes and distribution following a 240 mL dose of water in the fasted state.* Mol Pharm *2014;**11**:3039–47, http://pubs.acs.org/doi/abs/10.1021/mp500210c).*

stability-indicating assay should be used to determine the concentration of the drug substance in selected buffers. A drug substance is considered highly soluble when the highest dose strength is soluble in 250 mL or less of aqueous media over the pH range of 1–6.8; otherwise, the drug substance is considered poorly soluble. The volume estimate of 250 mL is derived from typical bioequivalence study protocols that prescribe administration of a drug product to fasting human volunteers with a glass (about 8 oz) of water.

12.3.1.2 Determination of drug substance permeability

The permeability classification can be determined in human subjects using mass balance or absolute bioavailability, which are the preferred methods, or intestinal perfusion approaches. The FDA draft guidance[15] cites the use of human pharmacokinetic studies and intestinal permeability methods as acceptable ways to determine permeability classification. In many cases, a single method may be sufficient. When a single method fails to conclusively demonstrate a permeability classification, two different methods may be advisable. In the absence of evidence suggesting instability in the GI, a drug substance is considered by the guidance to be highly permeable when the extent of absorption in humans is determined to be 85% or more of an administered dose based on a mass balance determination or in comparison to an intravenous reference dose.[15] Otherwise, the drug substance is considered to be poorly permeable.

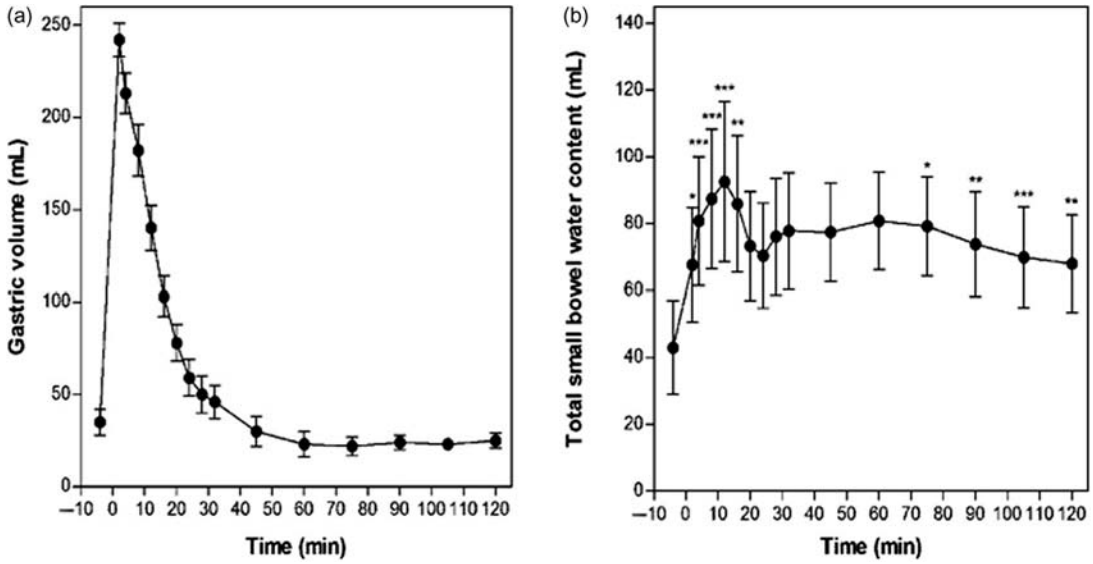


FIGURE 12.3 (a) Mean gastric volume and (b) mean total small bowel water content before and after ingestion of a 240-mL dose of water given at $t = 0$ min. $n = 12$ healthy volunteers. Error bars represent $\pm$SEM. Dunn's multiple comparison test versus baseline value $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$. *This figure reproduced with the permission of American Chemical Society (Mudie DM, Murray K, Hoad CL, Pritchard SE, Garnett MC, Amidon GL, et al. Quantification of gastrointestinal liquid volumes and distribution following a 240 mL dose of water in the fasted state.* Mol Pharm *2014;**11**:3039–47, http://pubs.acs.org/doi/abs/10.1021/mp500210c).*

12.3.1.2.1 Mass balance and absolute bioavailability studies

In mass balance studies, unlabeled stable isotopes or radiolabeled drug substances are used to determine the extent of absorption of a drug. A sufficient number of subjects should be enrolled to provide a reliable estimate of the extent of absorption. Oral bioavailability determination can be performed using intravenous administration as a reference. The obtained extent of absorption in humans based on mass balance studies using total radioactivity in urine does not take into consideration the extent of degradation of a drug in the GI fluid prior to intestinal membrane permeation. Therefore, a drug's stability in the GI tract needs to be investigated unless more than 85% of unchanged drug is excreted via the urine. Experimentally, drug solutions in simulated fluids such as gastric and intestinal fluid USP should be incubated at 37°C for a period that is representative of in vivo drug contact with these fluids (eg, 1 hour in gastric fluid and 3 hours in intestinal fluid). Over 5% degradation of a drug in this study could suggest potential instability.

12.3.1.2.2 Intestinal permeability

The 2015 draft guidance[15] recommends that the following methods be considered appropriate to determine the permeability of a drug substance from the GI: (1) in vivo human intestinal perfusion; (2) in vivo or in situ animal intestinal perfusion; (3) in vitro permeation studies using excised human or animal intestinal tissues; or (4) in vitro permeation across a monolayer of cultured epithelial cells. The advantages and disadvantages of each method in determination of intestinal permeability will be discussed later in this chapter. An apparent passive transport mechanism can be assumed when one of the following conditions is satisfied:

1. A linear pharmacokinetic relationship between the drug dose and its area under the concentration-time curve (AUC) is demonstrated in humans.
2. Independence of the measured in vivo or in situ permeability on initial drug concentration in the perfusion fluid is demonstrated in an animal model.
3. Independence of the measured in vitro permeability on initial drug concentration in donor fluid or on transport direction using a suitable in vitro cell culture method is demonstrated.

To demonstrate the suitability of a method intended for BCS-based permeability determination, a rank-order relationship between experimental permeability values and the extent of drug absorption data in human subjects should be established using a sufficient number of model drugs (Table 12.4).[15] The model drugs should represent a range of low (eg, <50%), moderate (eg, 50%–84%), and high (≥85%) absorption and should have available information on the mechanism of absorption and reliable estimates of the extent of drug absorption in humans.[15] For in vivo intestinal perfusion

TABLE 12.3 Biopharmaceutics Classification System

Biopharmaceutics class	Solubility	Permeability
I	High	High
II	Low	High
III	High	Low
IV	Low	Low

TABLE 12.4 The Model Drugs Suggested for Use in Establishing the Suitability of a Permeability Method

Model drugs in establishing suitability of a permeability method		
High permeability	**Moderate permeability**	**Low permeability**
($f_a \geq 0.85$)	**($0.50 \leq f_a \leq 0.84$)**	**($f_a < 0.50$)**
Antipyrine	Amiloride	Acyclovir
Caffeine	Atenolol	Chlorothiazide
Carbamazepine	Chlorpheniramine	Enalaprilat
Disopyramide	Creatinine	Famotidine
Ketoprofen	Enalapril	Foscarnet
Metoprolol	Furosemide	Lisinopril
Minoxidil	Hydrochlorothiazide	Mannitol
Naproxen	Metformin	Nadolol
Phenytoin	Ranitidine	Polyethylene glycol (400)
Propranolol	Terbutaline	Sulpiride
Theophylline		

studies in humans, six model drugs are recommended. For in vivo or in situ intestinal perfusion studies in animals and for in vitro cell culture methods, 20 model drugs are recommended.

After demonstrating the suitability of a method and maintaining the same study protocol, it is not necessary to retest all selected model drugs for subsequent studies intended to classify a drug substance. Instead, two internal standards of a low- and a high-permeability model drug, in addition to the fluid volume marker (or a zero-permeability compound such as PEG-4000), should be included in permeability studies. The permeability values of the two internal standards should not differ significantly between different tests, including those conducted to demonstrate suitability of the method. At the end of an in situ or in vitro test, the amount of drug in the membrane should be determined to assist in the calculation of mass balance.

12.3.1.3 *Comparison of dissolution profile*

In the BCS draft guidance,[15] an IR drug product is considered rapidly dissolving when over 85% of the labeled amount of the drug substance dissolves within 30 minutes, using USP Apparatus I at 100 rpm or Apparatus II at 50 rpm (or at 75 rpm when appropriately justified) in a volume of 500 mL or less in each of the following media: (1) 0.1 N of HCl or simulated gastric fluid without enzymes; (2) a pH 4.5 buffer; and (3) a pH 6.8 buffer or simulated intestinal fluid without enzymes. In addition, very rapid dissolution is defined to be over 85% of the labeled amount of the drug substance dissolved within 15 minutes using the abovementioned conditions.

According to the draft guidance,[15] observed in vivo differences in the rate and extent of drug absorption from two pharmaceutically equivalent solid oral products may be due to differences in in vivo drug dissolution. If the in vivo dissolution is rapid or very rapid in relation to gastric emptying, oral drug absorption is likely to be independent of drug dissolution. Thus, for BCS class I (high solubility/high permeability) or BCS class III (high solubility/low permeability) drug substances, so long as the inactive ingredients in the dosage form do not significantly affect the absorption of active ingredients, demonstration of in vivo bioavailability or bioequivalence may not be necessary for drug products containing class I or III drug substances.

12.3.2 Scientific basis for BCS

The key parameters controlling drug absorption are three dimensionless numbers: dose number, dissolution number, and absorption number, representing the fundamental processes of dose, drug dissolution, and membrane permeation, respectively.

The dose number (D_0) is characterized by the volume required for solubilizing the maximum dose strength of the drug[16]:

$$D_0 = \frac{M/V_0}{C_s} \tag{12.2}$$

where, C_s is the drug solubility, M is the dose, and V_0 is the volume of water taken with the dose, which is generally set to 250 mL.

The dissolution number (D_n) is characterized by the time required for drug dissolution, which is the ratio of the intestinal residence time (t_{res}) and the dissolution time (t_{diss})[16]:

$$D_n = \frac{t_{res}}{t_{diss}} = \frac{3DC_s}{r^2\rho} \times t_{res} \tag{12.3}$$

where D is diffusivity, ρ is density, and r is the initial particle radius.

The absorption number (A_n) is characterized by the time required for absorption of the dose administered, which is a ratio of residence time and absorptive time (t_{abs})[16]:

$$A_n = \frac{t_{res}}{t_{abs}} = \frac{P_{eff}}{R} \times t_{res} \tag{12.4}$$

where P_{eff} is the permeability and R is the gut radius.

Drugs with complete absorption show $D_0 < 1$, while D_n and $A_n > 1$. This approach is used to set up a theoretical basis for correlating in vitro drug product dissolution with in vivo absorption. To follow the movement of a dosage form through the GI, the entire process needs to be broken down into several component parts in order to see it mechanistically. The fundamental starting point is to apply Fick's first law to the absorption across the intestinal membrane:

$$J_w = P_w \times C_w \tag{12.5}$$

where J_w (x, y, z, t) is the drug flux (mass/area/time) through the intestinal wall at any position and time, P_w (x, y, z, t) is the permeability of intestinal membrane, and C_w (x, y, z, t) is the drug concentration at the intestinal membrane surface. This equation is a local law pertaining to each point along the intestinal membrane. It is assumed that sink conditions (ie, drug concentration equals zero) exist and that P_w is an effective permeability. The plasma may be assumed to be the physiological sink since concentrations in the plasma are generally several orders of magnitude lower than that in the intestinal lumen in humans. The drug absorption rate (ie, the rate of loss of drug from the intestinal lumen), assuming no luminal reactions, at any time t is

$$\text{Absorption rate} = \frac{dm}{dt} = \iint_A P_w C_w dA \tag{12.6}$$

where the double integral is over the entire GI surface.

The total mass, M, of drug absorbed at time t is

$$M_{(t)} = \int_0^t \iint_A P_w C_w dA dt \tag{12.7}$$

These mass balance relations are very general since the surface can be of arbitrary shape and the concentration at the membrane wall and permeability can have any dependence on position and time. For full generality, the permeability P_w must be considered to be position dependent as well as time dependent.

Based on Eqs. 12.2–12.7, the following principle for bioequivalence may be stated: if two drug products containing the same drug have the same concentration-time profile at the intestinal membrane surface, then they will have the same rate and extent of absorption. Furthermore, this statement may imply that if two drug products have the same in vivo dissolution profile under all luminal conditions, they will have the same rate and extent of drug absorption. These general principles assume that there are no other components in the formulation that affect the membrane permeability, intestinal transit, or both. Accordingly, the fraction absorbed (f_a) of a solution follows an exponential function, and can be calculated as follows:

$$f_a = 1 - e^{-2An} \tag{12.8}$$

As to the establishment of in vitro-in vivo correlation (IVIVC), several factors need to be considered. These in vitro dissolution tests can only model the release and dissolution rates of the drug, and it is only when these processes are rate limiting in the overall absorption that IVIVC can be established. For class I drugs, the complete dose will be dissolved already in the stomach and, provided that the absorption in the stomach is negligible, the gastric emptying will be rate-limiting and therefore IVIVC is not expected. Thus, in vitro dissolution testing can be expected to be overdiscriminating for those drugs, since tablets showing different in vitro dissolution profiles may provide the same rate and extent of bioavailability. Class II drugs are expected to have a dissolution-limited absorption and an IVIVC can be established using a well-designed in vitro dissolution test.[16] But the IVIVC will not be likely for class II drugs if absorption is limited by the saturation solubility in the GI tract rather than by the dissolution rate. In this situation, the drug concentration in the GI tract will be close to the saturation solubility and changes in the dissolution rate will not affect the plasma concentration profile or the in vivo bioavailability. The absorption of class III drugs is limited by their intestinal permeability and no IVIVC should be expected.[16] However, when the drug dissolution becomes slower than the gastric emptying, a reduction of the extent of bioavailability will be found at slower dissolution rates, since the time during which the drug is available for transport across the small intestinal barrier will then be reduced. The class IV drugs present significant problems for effective oral delivery. Very limited or no IVIVC is expected.

12.4 INTESTINAL PERMEABILITY EVALUATION: CULTURED CELLS

The term *permeability* is defined as the rate (cm/second) at which a solute is transported across the intestinal membrane, irrespective of transport mechanism. Various methodologies can be used for evaluation of intestinal permeability and drug absorption; each has its advantages and disadvantages. Numerous in vitro methods have been used in the drug selection process for

assessing the intestinal absorption of potential drug candidates with distinct advantages and drawbacks. Compared to in vivo drug absorption studies, evaluation of intestinal permeability in vitro requires less compound; is relatively easier and, in the case of regional absorption studies, avoids complicated surgery and maintenance of surgically prepared animals; is more rapid and has the potential to reduce animal usage since a number of variables can be examined in each experiment; provides insights into the mechanisms (eg, carrier-mediated vs passive), routes (eg, transcellular vs paracellular), and regional differences (eg, small vs large intestine) involved in transepithelial transport; and is analytically more simple because compounds are being analyzed in an aqueous buffer solution as opposed to whole blood or plasma. However, one universal limitation with all the in vitro systems is that the effects of physiological factors such as gastric-emptying rate, GI transit rate, and GI pH cannot be incorporated into the data interpretation.

The successful application of in vitro models to predict drug absorption across the intestinal mucosa depends on how closely the in vitro model mimics the characteristics of the in vivo intestinal epithelium. Although it is very difficult to develop a single in vitro system that can simulate all the conditions existing in the human intestine, combinational in vitro systems are used as decision-making tools in drug discovery and development. Various cell models that mimic in vivo intestinal epithelium in humans have been developed and widely used. The cell lines are now routinely cultivated as monolayers for studies of the transepithelial transport of drugs. The examples of some selected cell culture models are summarized in Table 12.5.

TABLE 12.5 Cultured Cells Commonly Used for Permeability Assessment

Cell type	Species or origin	Special characteristics
Caco-2	Human colon adenocarcinoma	Most well-established and frequently used cell model
		Differentiates and expresses some relevant efflux transporters
		Expression of influx transporters is variable (laboratory to laboratory
MDCK	Madin Darby canine kidney epithelial cells	Polarized cells with low intrinsic expression of ATP-binding cassette (ABC) transporters
		Ideal for transfections
LLC-PK1	Pig kidney epithelial cells	Polarized cells with low intrinsic transporter expression
		Ideal for transfections
2/4/A1	Rat fetal intestinal epithelia cell	Temperature-sensitive
		Ideal for paracellularly absorbed compounds (leakier pores)
TC-7	Caco-2 subclone	Similar to Caco-2
HT-29	Human colon	Contains mucus-producing goblet cells
IEC-18	Rat small intestine cell line	Provides a size-selective barrier for paracellularly transported compounds

12.4.1 Caco-2 cells

The Caco-2 cell line was first isolated in the 1970s from a human colon adenocarcinoma. During growth, Caco-2 cells go through processes of proliferation, confluency, and differentiation.[17] When grown under standard conditions on semipermeable membranes, fully differentiated Caco-2 cells are very similar to normal enterocytes with regards to their morphological characteristics. Primarily, they have functional tight junctions and they develop apical and basolateral domains and brush border cytoskeleton. Because the permeation characteristics of drugs across Caco-2 cell monolayers correlate with their human intestinal mucosa permeation characteristics, it has been suggested that Caco-2 cells can be used to predict the oral absorption of drugs in humans. In the past two decades, there has been a dramatic increase in the use of Caco-2 cells as a rapid in vitro screening tool in support of drug discovery within the pharmaceutical industry, as well as in advanced mechanistic studies, including intestinal absorption, transport, and metabolism[18,19]; utilization of absorption modifying excipient in the drug formulation[20]; and oral toxicity.[21]

The Caco-2 cells are usually grown in culture T-flasks (75 cm^2 or 175 cm^2) in a CO_2-incubator at 37°C with 5% CO_2. The culture medium is Dulbecco's Modified Eagle Medium (DMEM) supplemented with 0.1 mM of nonessential amino acids, 100 U/mL of penicillin, 0.1 g/mL of streptomycin, 10 mM of sodium bicarbonate, and 10% fetal bovine serum (FBS). At a confluency of 70%–80%, the cells are split at a ratio of 1:3–1:5. For use in the transport experiments, the cells are harvested with trypsin-EDTA and seeded on polycarbonate filters (0.4 μm pore size, 1.13 cm^2 growth area for 12-well plate, 0.33 cm^2 growth area for 24-well plate) inside Transwell cell culture chambers at a density of approximately 75,000 cells/cm^2 (Fig. 12.4). The culture medium is refreshed every 24–48 hours. Transport studies are usually done after 18–21 days in culture, when expression of transporters (eg, P-glycoprotein (P-gp)) reaches its maximum.

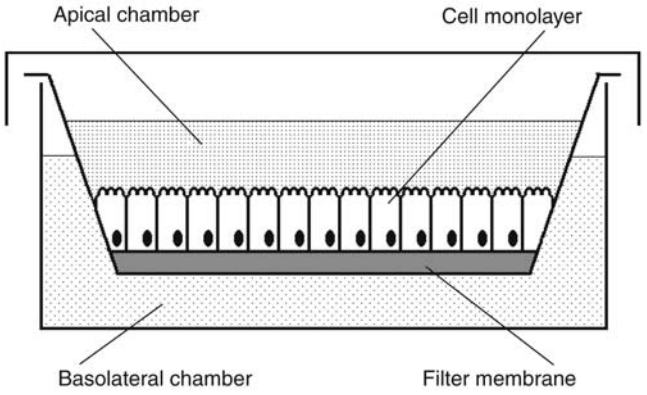


FIGURE 12.4 Typical experimental setup of a Caco-2 system used to study the transport of molecules across a monolayer.

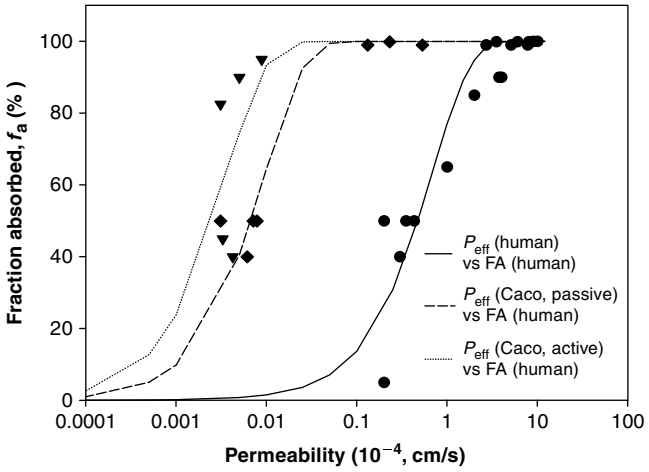


FIGURE 12.5 The relationship between human jejunal effective permeability, Caco-2 monolayer permeability, and the fraction dose absorbed in humans.

To ensure the integrity of the monolayer during the course of the experiment, quality control is done by measuring the transepithelial electrical resistance (TEER) values and the permeability of mannitol (low cellular permeability). Caco-2 inserts with TEER values greater than $250\,\Omega \times cm^2$ (24 wells) can be used in the study. The permeability of ^{14}C- or ^{3}H-mannitol should be lower than 0.5×10^{-6} cm/second. In each experiment, certain reference compounds (mannitol and propranolol with high cellular permeability) should be included. The apparent permeability (P_{app}) coefficient (expressed in cm/second) is calculated as follows:

$$P_{app} = \frac{V_R}{A \times C_0} \times \frac{dC}{dt} \tag{12.9}$$

where, V_R is the volume in the receiver chamber (mL), A is the filter surface area (cm^2), C_0 is the initial donor concentration of the drug (μg/mL), and dC/dt is the initial slope of the cumulative concentration (μg/mL) in the receiver chamber with time (second). If the P_{app} values in the apical-to-basolateral direction are not equal to those in the basolateral-to-apical direction, or if the P_{app} values decrease with increasing test compound concentration, this suggests the involvement of transporter pathways. Additional incubations may then be carried out to probe for the involvement of specific transporter or efflux systems. Completely absorbed drugs are found to have a high permeability coefficient ($P_{app} > 1 \times 10^{-6}$ cm/second), whereas incompletely absorbed drugs generally have a low permeability coefficient ($P_{app} < 1 \times 10^{-7}$ cm/second) in Caco-2 cells. Caco-2 cells have been demonstrated as an excellent model of the passive transcellular pathway,[22] or model comparison with parallel artificial membrane permeation assay (PAMPA).[23] Certain studies suggest that Caco-2 cells rank the permeability of drugs in the same order as more complex absorption models such as in situ perfusion models. These correlation studies are mainly performed with passively transported drugs (Fig. 12.5).

12.4.2 Limitations of Caco-2 cell model

There are a few important factors that limit the use of cultured cell models. The lack of correlation for paracellularly transported compounds is one limitation.[22] Caco-2 has been shown to have a significantly lower number of paracellular pores than the human intestine.[22] Low-molecular-weight hydrophilic compounds (eg, ranitidine, atenolol, furosemide, hydrochlorothiazide) displayed low permeability in the Caco-2 cell model despite the extent of absorption in humans is greater than 50%. In other words, this cell model can only serve as a one-way screen such that compounds with high permeability in this model are typically well absorbed in vivo; however, compounds with low permeability cannot be certainly classified as poorly absorbed compounds in vivo.

Another limitation is that the carrier-mediated drug intestinal absorption is poorly predicted with the Caco-2 cell model.[22] The Caco-2 cell model primarily measures passive drug transport (both transcellular and paracellular). Caco-2 cells can express some transporters (eg, peptide transporters, organic cation transporters (OCTs), and organic anion transporters (OAT)), but most intestinal transporters are quantitatively underexpressed when compared to those in vivo.[22,24] Thus, this model is likely to generate false negatives with actively transported drug candidates.

In addition, the correlation of the permeability of slowly and incompletely absorbed drugs in the Caco-2 cells and human jejunum seems qualitative rather than quantitative. These drugs (eg, lamivudine)[24] are

transported at a 30- to 80-fold slower rate in the Caco-2 cells than in the human jejunum. It is also very difficult to compare the absolute permeability coefficient value of individual compounds reported in the literature, particularly with compounds that are primarily permeate via the paracellular route. The variability may be attributed to differences in culture conditions and composition of cell subpopulation.

Additionally, there is the lack of metabolic enzymes in cell models. An ideal cell-based intestinal permeability tool would be one that simulates the human GI enterocytes not only in lipid bilayer characteristics, but also in metabolic enzyme activity. Although Caco-2 cell models are known to express adequate amounts of hydrolase, esterase, and brush-border enzymes, they fail to simulate the complete in vivo intestinal environment because they do not express appreciable quantities of cytochrome P450 (CYP) 3A4, the principal CYP present in human enterocytes.

Moreover, the use of appreciable amount of organic cosolvent is limited. The integrity of tight junctions is easily compromised by commonly used organic solvents (eg, methanol, ethanol, PG, PEG) even at a small concentration (more than 1%–2% v/v). Therefore, a significant percentage of new drug candidates with poor aqueous solubility cannot be evaluated in this model. Furthermore, cultured cell systems cannot be used to adequately evaluate those compounds that can nonspecifically bind to plastic devices.

In most cases, the preparation of fully functional cell monolayers generally requires a 3-week cell culture period with 8–9 instances of laborious cell feeding. However, the preparation time can be substantially reduced by modifying both the coating material and growth media. A shorter cell-culturing period to generate functional monolayers not only increases the overall productivity, but also reduces the chance of bacterial/fungal contamination so as to minimize downtime.

12.4.3 MDCK cells

The Madin-Darby canine kidney (MDCK) line, established in the 1950s, is another of the most frequently used cell lines for permeability assessment. MDCK cells are grown in the culture medium Modified Eagle Medium (MEM) supplemented with 2% sodium bicarbonate, 1% glutamine, and 5% FBS. Like Caco-2 cells, MDCK cells have been shown to differentiate into columnar epithelia and to form tight junctions when cultured on semipermeable membranes. Given the fact that Caco-2 cells are derived from human colon carcinoma cells, whereas MDCK cells are derived from dog kidney cells, the major difference between Caco-2 and MDCK cells is the absence or presence of active transporters and drug metabolizing enzymes. The use of MDCK cells has been demonstrated as a tool for assessing the membrane permeability properties of early drug discovery compounds.[25,26]

For passively transported compounds, MDCK cells can be used as an alternative for Caco-2 cells.[26–28] Although a high rank correlation coefficient of 0.93 for MDCK to Caco-2 P_{app} has been observed, the relationship between the permeability in cells and the absorption in humans is less strongly correlated in MDCK. Permeability values tend to be lower in MDCK cells than in Caco-2 cells. While for actively transported compounds, the discrepancy is more significant between these two cell models. One major advantage of MDCK cells over Caco-2 cells is the shorter cultivation period (3 days vs 3 weeks). A shorter cell culture time becomes a significant advantage considering reduced labor and reduced downtime in case of cell contamination. However, species differences should also be considered before using MDCK cells as a primary screening tool for permeability in early drug discovery.

12.4.4 Other cells

TC-7 is one of the subclones isolated from Caco-2 cells with very similar cell morphology, concerning apical brush border, microvilli, tight junctions, and polarization of the cell line. The dipeptide transporter is underexpressed in both cell lines. The TC-7 clone, however, appears more homogenous in terms of cell size. Both cell lines achieve similar monolayer integrity toward mannitol and PEG-4000. A comparison in apparent drug permeability has been made between TC-7 cell and its parental Caco-2 cells.[29] When using cyclosporine A as a P-glycoprotein substrate, active efflux is shown to be lower in the TC-7 clone than in the parental Caco-2 cells. The correlation curves obtained with both cell types are almost completely superimposable, based on the permeability result from 20 test compounds exhibiting large differences in chemical structure, molecular weight, transport mechanisms, and percentage of absorption in humans.

Lewis lung carcinoma-porcine kidney 1 (LLC-PK1) cells have also been explored as an alternative to Caco-2 cells for assessing the permeability of test compounds.[30] Several investigators have reported the utility of porcine cell line for characterizing the passive absorption discovery compounds.[31]

The cell model of 2/4/A1, which originates from fetal rat intestine, is found to be able to better

mimic the permeability of the human small intestine, particularly with regard to passive transcellular and paracellular permeability.[32,33] This immortalized cell line forms viable differentiated monolayers with tight junctions, brush-border membrane enzymes, and transporter proteins. The paracellular pore radius of 2/4/A1 cell has been determined to be 9.0 ± 0.2 Å, which is similar to the pores in the human small intestine, whereas the pore size in Caco-2 is much smaller (about 3.7 ± 0.1 Å). Therefore, 2/4/1A cells have been proposed as a better model to study passively transported compounds via the paracellular route than Caco-2 cells. The transport rate of poorly permeable compounds (eg, mannitol and creatinine) in 2/4/1A monolayers is comparable with that in the human jejunum and is up to 300 times faster than that in the Caco-2 monolayers, suggesting that the 2/4/1A cell line will be more predictive for compounds that are absorbed via paracellular route.

12.5 INTESTINAL PERMEABILITY EVALUATION: EX VIVO

12.5.1 The everted gut sac technique

The everted sac approach introduced in 1954 by Wilson and Wiseman,[34] has been widely used to study intestinal drug transport. Everted sacs are prepared from rat intestine by quickly removing the intestine from the decapitated animal, flushing with a saline solution, everting it over a glass rod (diameter 3 mm), filling it with fresh oxygenated culture medium, and dividing it into sacs approximately 2–4 cm in length with silk suture (Fig. 12.6). Sacs (with the mucosal side outside) are then submerged in a culture medium containing the drug of interest, and accumulation in the inner compartment is measured. Under optimal conditions, sacs remain viable for up to 120 minutes. To check sac viability and integrity, glucose concentration can be monitored inside and outside the everted sac during the experiments. It is a two-compartment system permitting the calculation of the kinetics of uptake into the epithelial tissue by assaying material in the cells, and transferring across the entire epithelial layer by measuring material that is found inside the sac at the end of the incubation period. The system has been used to study the uptake of liposomes and proteins, bioadhesive lectins, and synthetic, nondegradable polymers.[35]

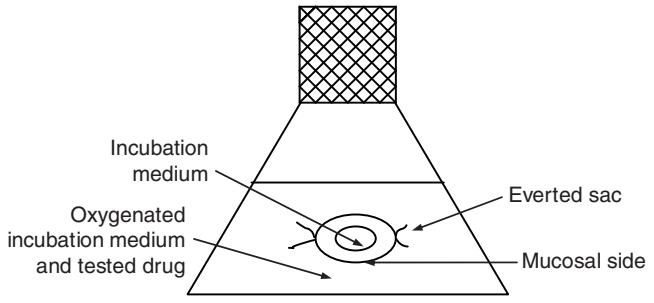


FIGURE 12.6 Everted sac: the sac is incubated by shaking in the oxygenate tissue culture medium. Each sac is removed at the required time point and drug is analyzed in the serosal space and tissue.

The everted sac technique is a useful tool to study the mechanisms of uptake of molecules. Test compounds are incubated with the sac at various concentrations and for different time periods to obtain the kinetics of uptake. At the end of the incubation, the compounds are analyzed in the tissue and the serosal space, and as tissue protein is measured, the uptake is expressed per unit protein for normalization. In the case of passive membrane diffusion, the compound is detected in both tissue and serosal space, and uptake will be concentration dependent. For carrier-mediated transports, uptake will be saturable and subject to competition by other substances transported by the carriers. Test compounds should be detected in both tissue and serosal space. In the case of active mechanisms, there will be inhibition by the metabolic inhibitors added to the system. Endocytosis is slow and may or may not be saturable, depending on whether it is receptor mediated, and inhibited by both metabolic inhibitors and inhibitors of microtubule assembly like colchicine. The compound will be detected in the tissue, and in very small amounts in the serosal space. If it is transported by the paracellular route, there will be no test compound in the tissue, only in the serosal space, and translocation should be concentration dependent. In addition, disruption of the tight junctions with the Ca^{2+} chelators ethylene diamine tetraacetic acid (EDTA) or ethyleneglycol tetraacetic acid (EGTA) should increase the uptake of molecules by the paracellular route.

Given the presence of P-gp activity along the rat intestine, the everted gut sac model can also be useful to study the action of intestinal P-gp on intestinal drug absorption. By comparing the transport kinetics of a substrate in the absence or the presence of potential P-gp inhibitors, the method can be used to evaluate the role of P-gp in the intestinal drug absorption and to screen for putative P-gp inhibitors. An understanding of the site specificity of absorption is most important to optimizing oral drug delivery, particularly in developing controlled released formulations.

The everted sac system affords a simple technique to investigate differences in drug absorption along the GI tract. The everted sac system is also a useful tool for screening absorption enhancers. The transport mechanism, which is modified by the enhancers, can be verified by using specific model compounds to measure the absorption in the presence and absence of enhancers. For example, quinine or digoxin for membrane diffusion, amino acids, or dipeptides for carrier transport, as well as mannitol for the paracellular route, can be used for this purpose. The everted gut sac has the potential for studying the metabolism of drugs by the intestinal mucosa. If drugs are metabolized by cell surface enzymes in the intestine, then the metabolic products will be detected in the medium following incubation. If the drug is absorbed into the cells, then it may be susceptible to first-pass metabolism within the enterocytes. The products of this metabolism may themselves be transported, in which case they would be detected in the serosal space, or they may be confined to the interior of the cells.

Thus, the everted gut sac system has a number of advantages. It is simple, quick, very reproducible, and inexpensive. The regional differences in drug absorption can be studied and differentiated with this system. It can provide information on the mechanism of drug absorption and be used to test the effects of enhancers and formulations on absorption. The system also has some disadvantages. The major disadvantage is that it is an animal model, and the absorption results obtained with this model may not always reflect the real absorption profile in humans. In addition, the test compounds have to across all the layers (including muscle) of the small intestine, instead of just the intestinal mucosa. Furthermore, the sink condition cannot be established due to the small volume inside the sac.

12.5.2 Ussing chamber

The Ussing chamber technique resulted from the pioneering work of Ussing and coworkers, who in the late 1940s and early 1950s published a series of papers describing the transepithelial ion fluxes. This technique has subsequently been modified by Grass and Sweetana to study drug transport. This technique utilizes small intestinal sheets that are mounted between two compartments. The mucosal and serosal compartments are usually supplied with Krebs-Ringer bicarbonate buffer (KRB), which is continuously gassed with a mixture of O_2:CO_2 (95:5). The test compound is added to either the mucosal or serosal side of the tissue to study transport in the absorptive or secretory direction, respectively. Apparent permeability coefficients (P_{app}) are calculated as for Caco-2 transport experiments from the appearance rate of the compound in the receiver compartment.

The applications of the Ussing chamber system include studying the transepithelial drug transport in combination with intestinal metabolism and investigating the regional differences in intestinal absorption.[36] Advantages of the technique are that the amount of drug needed to perform a study is relatively small, and the collected samples are analytically clean. Unlike in vitro cultured cell model, the presence of the apical mucus layer makes the system more relevant to in vivo condition. A novel factor called the *permeability ratio (PR)*, which is the value of in vitro intestinal permeability obtained from the Ussing chamber, further corrected by the permeability of lucifer yellow, a paracellular permeability marker, was recently proposed for reliable in vitro permeability.[37] By using this approach, the correlation with human in vivo intestinal permeability of 12 tested model drugs representing different BCS classes is high.[37] More important, the new method perfectly classifies all 12 model drugs.

In addition, the model offers the possibility to study bidirectional drug transport and concentration dependency of transport and the effect of specific transport inhibitors. The Ussing chamber model has been used to demonstrate transport polarity of compounds that are substrates of P-gp or the multidrug resistance (MDR)−associated protein.[38] Furthermore, species differences with respect to intestinal absorption characteristics can be determined, which can be useful during the selection of a suitable animal model for drug bioavailability studies. It should be noted that during preparation of the Ussing chamber system, dissection of the epithelial tissue is rather difficult to perform and the serosal muscle layers can be only partially removed, which may result in an underestimation of transport, which appears to be an issue, especially for lipophilic drugs.

12.5.3 In situ intestinal perfusion in rat

Among the various preclinical models that are used to study drug absorption, the in situ method is the nearest to the in vivo system. Intestinal segments of the anesthetized animals are cannulated and perfused by a solution of the drug and the amount of the drug that is taken up from the perfusate can be calculated (Fig. 12.7). Input of the drug compound can be closely controlled in terms of concentration, pH, osmolality,

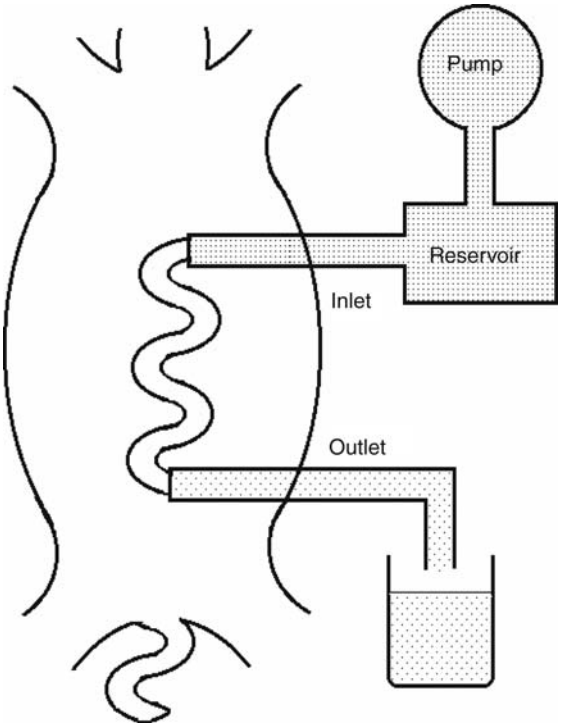


FIGURE 12.7 The in situ study of drug uptake by the single-pass perfusion technique. A rat is anesthetized, and an intestinal segment is cannulated and perfused with the solution containing the compound of interest. The amount of the compound taken up from the solution can be calculated from the difference between the inlet and the outlet concentration of the compound.

intestinal region, and flow rate. The advantages of in situ method include:

1. The barriers that a compound has to cross to reach the portal blood circulation are identical in the in situ and in vivo situations.
2. This method provides a unique possibility for studying the intestinal events under isolated status without the complication of biliary excretion and enterohepatic circulation.
3. It is good for producing some kinetic data, particularly as the blood supply, innervation, and clearance capabilities of the animal remain intact.
4. Like the Ussing chamber technique, the in situ method can also be used for the regional absorption studies, and observing the gross effects of enhancer. Thus, the method has been widely used for the selection of drug candidates. And in situ intestinal perfusion in rat jejunum in the predictive accuracy of uptake and efflux transport seems superior to Caco-2 cell model.[24]

The prediction of the extent of absorption in humans from rat intestinal perfusion data has been demonstrated for a series of compounds with variable absorption (5%−100%) in humans. A high correlation between effective permeability values determined in rat and human jejunum has been demonstrated. Therefore, in situ perfusion of rat jejunum is considered to be a useful tool to classify compounds according to BCS,[39,40] provided that appropriate reference compounds are included to account for interlaboratory variations in passively and actively transported drugs.

The basic experimental procedure is as follows. Rats are fasted for at least 12 hours and anesthetized, and then an intestinal segment (10−20 cm) is isolated and cannulated. This segment is rinsed with an isotonic solution before being constantly perfused with the solution containing the compound of interest. The perfusate is then collected in determined intervals. Complete recovery of PEG-4000 (a nonabsorbable marker), stable water flux, and the effective permeability coefficient (P_{eff}), with time for compounds transported both passively (antipyrine) and by a carrier-mediated mechanism (glucose), indicate that viability of the intestinal tissue is maintained during the experiment. The P_{eff} of antipyrine can also be used as an indication of extensive change of the mesenteric blood flow. P_{eff} is calculated as follows:

$$P_{eff} = \frac{-Q_{in} \times \ln(C_{out}/C_{in})}{2\pi rL} \qquad (12.10)$$

where C_{in} and C_{out} are the inlet and outlet concentration of the compound in the perfusate, Q_{in} is the flow rate of perfusion medium entering in the intestinal segment, and $2\pi rL$ is the mass transfer surface area within the intestinal segment, which is assumed to be the area of a cylinder with length L and radius r. The accessibility of the luminal compartment enables an estimation of intestinal metabolism to be made. However, perfusion systems give no information about events at the cellular or membrane level. The measurement of the rate of decrease of the drug concentration in the perfusate does not always represent the rate of absorption of the drug into the systemic circulation, particularly if presystemic or intracellular intestinal metabolism occurs. This problem can be overcome if this technique is applied in combination with measurements of the concentration of the drug in the portal vein. The method is also limited because of its cost factor, as it requires a large number of animals to give statistically significant absorption data. Relatively high amounts of test compounds are also required to perform studies, which is not feasible in early drug discovery.

Also, surgical manipulation of the intestine combined with anesthesia can cause a significant change in the blood flow to the intestine and may have a remarkable effect on absorption rate. More often, the method uses perfused intestinal flow rates higher than those in vivo; and as a consequence, the in situ intestine is fully distended and the luminal hydrostatic pressure increases. This distension may affect the intestinal permeability or absorptive clearance.

12.5.4 Intestinal perfusion in humans

Intestinal perfusion techniques have been first used in humans with the aim of studying the absorption and secretion mechanisms or the role of gastric emptying in drug absorption, while the techniques are too complex to be used for mass or routine screening studies. The basic principle of these techniques is to infuse a solution of the test compound and a nonabsorbable marker such as PEG-4000 into an intestinal segment and to collect perfusate samples after the perfusion has passed through the segment. The difference between the inlet and the outlet concentration of the compound in the solution is assumed to have been absorbed.

Three main single-pass perfusion approaches have been employed in the small intestine:

- Open system—A triple lumen tube including a mixing segment
- Semiopen system—A multilumen tube with a proximal occluding balloon
- Closed system—A multilumen tube with two balloons occluding a 10 cm long intestinal segment (Fig. 12.8).[41]

In the open system method, entering perfusion solution and GI fluids are mixed in a "mixing segment," and at the distal end of the mixing segment, a sample is taken that is considered to be the inlet concentration of the test segment. The absorption is then calculated from a second outlet sample taken at the end of the test segment, which usually is 20−30 cm distal to the mixing segment. A major disadvantage of this method is that the composition of the perfusate will change along both mixing and test segments, which makes it difficult to define the absorption conditions and therefore to determine reference drug permeability at well-controlled luminal conditions. Perfusate can flow in either direction, and it is difficult to estimate the actual segment length in this open system. The semiopen system, which has been introduced by Ewe and Summerskill, overcomes the problem of proximal contamination by using an occluding balloon proximal to the test segment. This method decreases proximal leakage, and therefore, the luminal composition will be maintained at equilibrium and drug permeability can be determined under well-controlled conditions. However, both the open and semiopen methods generally use rather high perfusion flow rates (typically between 5 and 20 mL/minute), which are significantly higher than physiological flow rates of 1−3 mL/minute. The recovery of liquid is low and variable with the length of the intestinal segment studied unknown. Lennernas et al. refined the technique by developing a new intestinal perfusion instrument, which consists of a multichannel sterile tube with two inflatable balloons, creating a 10-cm-long segment and allowing segmental intestinal perfusion (Fig. 12.8).[42] The tube is introduced orally after local anesthesia, and when the tube has been positioned, the balloons are inflated, creating a closed intestinal segment. The segment is rinsed, perfused with a drug-free perfusate and then perfused with the perfusion solution containing the test compound. The absorption rate is calculated from the disappearance rate of the compound from the perfused segment. The absorption rate can be calculated in different ways, but the intestinal P_{eff} most likely provides the best description of the transport process across the intestinal barrier. The P_{eff} is calculated as shown in Eq. 12.10. There is a good correlation between the measured effective permeability values and the extent of absorption of drug in humans determined by pharmacokinetics studies (Fig. 12.9).

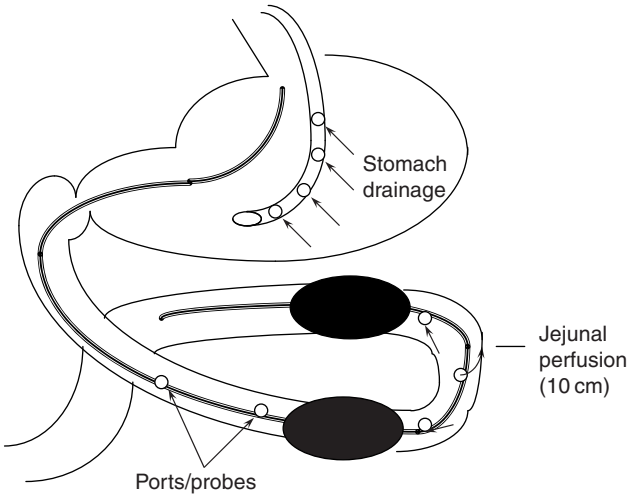


FIGURE 12.8 The multichannel tube system with double balloons enabling segmental perfusion in humans. The difference between the inlet and the outlet concentration of the compound in the solution is assumed to have been absorbed.

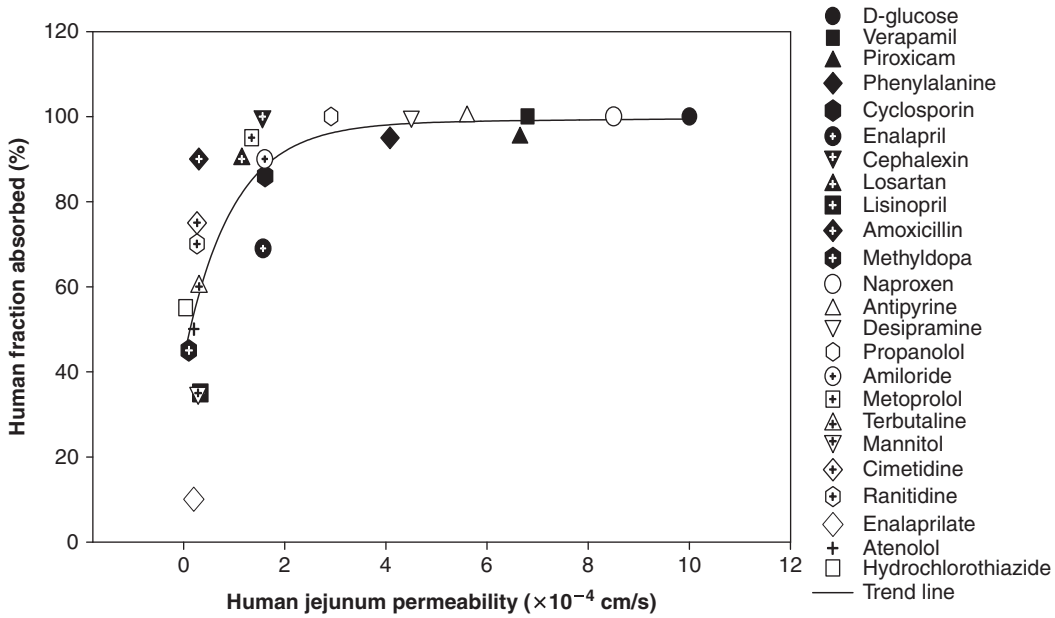


FIGURE 12.9 The relationship between human jejunal effective permeability and fraction dose absorbed.

12.6 IN SILICO METHODS

Computational screening has received much attention in the last few years. In silico models that can accurately predict the membrane permeability of test drugs based on lipophilicity, H bonding capacity, molecular size, polar surface area (PSA), and quantum properties has the potential to specifically direct the chemical synthesis and therefore revolutionize the drug discovery process. Lipinski's Rules of Five predict lower developability for compounds with more than 5 H bond donors, 10 H bond acceptors, molecular weight greater than 500, or *c* log *P* greater than 5.[43] Using this empirical model, useful predictions were achieved for closely related analog series of compounds. Additionally, approaches of the quantitative structure-activity relationship (QSAR) model and quantitative structure-property relationship (QSPR) model, and physiologically based pharmacokinetic (PBPK) modeling such as the compartmental absorption and transit (CAT) model have been developed for predicting oral drug absorption and bioavailability in humans.

12.6.1 QSAR

In last decade, significant progress has been made in the utilization of the QSAR approach in the early stages of drug discovery or development. The quantitative structure–human effective intestinal membrane permeability (P_{eff}) relationship was constructed based on seven structural parameters.[44] The predicted P_{eff} values had a good correlation with the experimental human absorbed fraction (F_a). This model was also used to perform high/low P_{eff} classification for 57 drugs that have been classified according to the BCS, and 72% of drugs could be classified correctly.

Several studies provided comparative QSAR analysis on PAMPA/modified PAMPA permeability.[45–47] One study predicted the PAMPA permeability with in silico descriptors, such as lipophilic partition coefficients (Log *P*), pK(a), and PSA.[45] The PAMPA prediction equation obtained in both studies could be applied for the prediction of both Caco-2 cell permeability and human intestinal absorption of mainly passively transported drugs.[45,46]

Prediction and mechanistic interpretation of human oral drug intestinal absorption using QSAR analysis in the stage of drug discovery and lead optimization has also been reported.[47–49] The important descriptors in oral absorption include log *P*, hydrogen bonding, and structural symmetry.[49] The final validated models are consistent and robust for the consensus prediction of oral drug intestinal absorption and are suitable for virtual screening applications.[47]

12.6.2 QSPR

Significant progress has been made in the utilization of this approach in the early stage of drug discovery. One study presented new linear and nonlinear predictive QSPR models to predict the human intestinal absorption rate.[50] The structure-property relationships so obtained involve only four molecular descriptors

and display an excellent ratio of number of cases to number of descriptors. The QSPR model has been reported for estimating both $\log P_{app}$ and the human intestinal absorption.[51] The results showed that the multivariate-statistics approach performs the best for estimating P_{app}, while a single variable like PSA describes the human intestinal absorption process satisfactorily.

12.6.3 PBPK modeling

Again, significant enhancement in the ability of PBPK models to successfully predict oral drug absorption and advance their role in preclinical and clinical development, as well as for regulatory applications, has been observed.[52,53] Predictive absorption models are used to determine the rate and extent of oral drug absorption, facilitate lead drug candidate selection, establish formulation development strategy, and support the development of regulatory policies.[54] The compartmental models mainly include the CAT model; advanced compartmental absorption and transit (ACAT) model; and advanced dissolution, absorption, and metabolism (ADAM) model.[54] The initial CAT model is a dynamic, mechanistic model that simulates and predicts the rate and extent of drug absorption from the human small intestinal tract. The CAT model has been modified and extended to be the ACAT model, which laid foundation for the commercial software, GastroPlus™.[55] More recently, the ADAM model was also developed based on the CAT model to include drug dissolution, GI fluid transit, drug degradation, gut wall permeation, intestinal metabolism, and active transport.[55] The Simcyp® software employs the ADAM model. Details of the prediction of oral drug absorption by PBPK modeling and simulation approach are described in chapter "Predictive Biopharmaceutics and Pharmacokinetics: Modeling and Simulation" of this book.

12.7 IN VIVO METHODS TO DETERMINE ORAL DRUG ABSORPTION

12.7.1 Mass balance study to determine drug absorption

Mass balance using radiolabeled drugs can be useful to estimate the fraction of dose absorbed. ^{14}C is the most frequently used radioisotope to label a drug, in nearly half of all studies, while 20% of studies are performed with ^{3}H. Mass balance studies to determine drug oral absorption have been available for certain drugs in animals[56] or humans.[57,58] Mass balance is determined based on the comparison of the radioactivity in the original dose to the amount of radioactivity in the excreta. Radioactivity measurements are independent of chemical structure, so total radioactivity measurements are the summation from the parent drug and its metabolites. The major advantage of mass balance study using radiolabeled drugs is that the total concentration of the parent drug and its metabolites can be quantified rather easily in various biological samples by determining the total radioactivity. Otherwise, the respective reference compound for each single metabolite will be needed. For animals, urine, feces, and bile samples dosed with the radiolabeled drug in question are collected over time, and the total radioactivity (of the drug and its metabolites) in the urine and bile sample reflects the actual amount of the drug absorbed into the GI enterocytes. For humans, normally, only urine and feces are collected; radioactivity in the urine suggests that the bioavailability is at least equal to the fraction of drug appearing in the urine; the amount of drug in the feces is a mixture of the unabsorbed drug (assuming oral administration), drug excreted in the GI tract, and drug excreted in the bile.

12.7.2 Rate of oral drug absorption into systemic circulation

After oral administration, the rate of the change in the drug amount in the body depends on the difference between the rate of drug absorption (rate of drug input) and the rate of drug elimination (rate of drug output):

$$\frac{dD_B}{dt} = \frac{dD_{GI}}{dt} - \frac{dD_E}{dt} \tag{12.11}$$

where D_B is the drug amount in the body, D_{GI} is the drug amount in the GI tract, and D_E is the drug amount eliminated. This equation is independent of the zero-order or first-order rate. Plasma drug concentrations over time in in vivo studies are usually measured.

The obtained plasma concentration-time curve for a drug following oral administration often presents several phases, including the absorption phase, time to reach peak concentration, postabsorption phase, and the terminal elimination phase (Fig. 12.10). During the absorption phase of a drug plasma concentration-time curve, the rate of drug input is greater than the rate of drug output; then, at the moment of drug peak plasma concentration, the rate of drug input equals the rate of drug output. Further, during the postabsorption phase of a drug plasma concentration-time curve, the rate of drug input is less than the rate of drug output; eventually, during the elimination phase when the drug at the absorption site becomes depleted, the rate

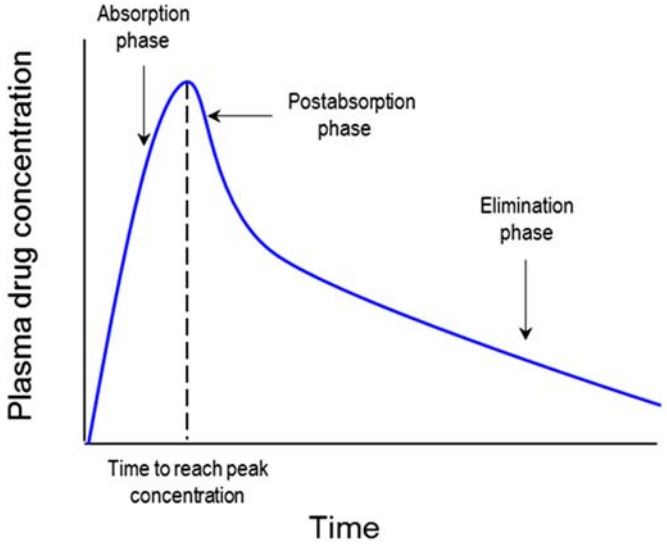


FIGURE 12.10 A typical plasma concentration-time curve for a drug after oral administration.

of drug input approaches 0, the rate of the change in the drug amount in the body represents only drug elimination from the body.

In pharmacokinetics, the overall rate of drug absorption could be described as either a first-order, or zero-order, or a mixed-input process. The drug absorption rate represents the net result of processes of drug dissolution, GI motility, blood flow, and drug transport across the membrane and into systemic circulation. Knowledge of drug absorption and drug elimination allows the prediction of steady-state drug concentration profiles. Analysis of drug absorption with the information of in vivo peak drug concentrations or time to achieve peak concentrations is also useful for the evaluation of the bioequivalence of generic drugs.

12.7.2.1 *First-order drug absorption*

Most pharmacokinetic models assume first-order kinetics for oral absorption. First-order drug absorption applies mostly to the oral absorption of drugs in solution or rapidly dissolving drug dosage forms in tablets or capsules (IR). Only drug in solution is absorbed into the body. The rate of drug input or the rate of drug disappearance from the GI is described by

$$\frac{dD_{GI}}{dt} = K_a D_{GI} F \tag{12.12}$$

where K_a is the first-order drug absorption rate constant from the GI, D_{GI} is the drug amount in solution in GI at a given time t, and F is the bioavailability. There are a few methods to calculate K_a based on in vivo data—namely, the method of residuals, mass balance methods (Wagner-Nelson method, Loo-Riegelman method), and the deconvolution method.[59]

12.7.2.2 *Zero-order drug absorption*

Zero-order drug absorption can also occur with oral drug absorption by a saturable process or a controlled release drug delivery system. Following oral drug administration, by zero-order drug absorption, drug in the GI is absorbed systemically at a constant rate of K_{ao}. The rate of drug input is simply K_{ao}:

$$\frac{dD_{GI}}{dt} = K_{ao} \tag{12.13}$$

where K_{ao} is the rate of drug absorption, which is a constant. In this case, the time for complete drug absorption to occur is equal to Dose of the drug/K_{ao}.

12.8 FOOD EFFECTS ON DRUG INTESTINAL ABSORPTION

Food-drug interactions in the GI have been widely associated with alterations of pharmacokinetic and pharmacodynamic parameters and proven to have significant clinical implications. Ideally, it is most advantageous if an oral drug administration can be provided independent of meal considerations. In the early stages of drug development, the food-drug interaction model would be beneficial when preclinical predictions could be of particular use and service to the industry. Although various in vitro and in vivo models can be found, no standard system currently exists to predict drug absorption and the effect of food.

12.8.1 GI physiological changes under fed state

Food effects on bioavailability are generally greatest when the drug product is administered shortly after a meal is ingested. The role of food in oral drug absorption may be attributed to a myriad of variables, ranging from the drug's physicochemical properties to the postprandial environmental changes in the GI. The environment of the GI varies markedly following meal ingestion. Food can alter a drug's transit time, luminal dissolution, drug permeability, and systemic availability by various means, including:

1. Increased gastric secretion and changed properties of gastric fluids (Tables 12.1 and 12.2); during digestion, gastric juice is secreted by the stomach wall into the lumen in order to provide various functions, including protein and lipid digestion, preparation of micronutrients for absorption in duodenum, protection against microbial overgrowth, dilution of chyme, and, relatedly, the

regulation of gastric emptying. The daily production of gastric juice ranges from 2−3 L with differences between the unstimulated and stimulated state. The gastric juice production under stimulation can be 10−50 mL/min, versus the unstimulated secretion of 1 mL/minute.[60]
2. Changed GI pH and viscosity; the nature of the gastric content is the mixed result of salivary flow, gastric secretions, and food properties.
3. Delayed gastric emptying and motility; generally, after a meal, the stomach is able to gather considerably more than 1 L of gastric content, whereas this maximum capacity depending on food volume and individual physiology often results in the gastric volume being large enough to exceed the gastric emptying. Due to long gastric transit times, the postprandial stomach plays an essential role in drug release and the appearance of food effects.
4. Stimulated bile flow; the levels of phospholipids and bile salts in the gut may increase fourfold to fivefold following a meal.[61]
5. Increased splanchnic blood flow.
6. Changed luminal metabolism (presystemic metabolism) of a drug substance; it has been well recognized that intake of grapefruit juice can markedly augment oral drug bioavailability (ie, dihydropyridines, terfenadine, saquinavir, cyclosporin, midazolam, triazolam, and verapamil and may also occur with lovastatin, cisapride, and astemizole), predominantly by inhibition of CYP3A4, and presumably additional inhibition of P-glycoprotein in the intestinal wall.[62] Since the duration of inhibitory effect of grapefruit juice can last 24 hours, repeated juice consumption can result in a cumulative increase in drug exposure.[63]
7. Physical or chemical interaction with a dosage form or a drug substance; in general, meals that are high in total calories and fat are more likely to affect the GI tract physiology, thereby resulting in a larger effect on the bioavailability of a drug substance or drug product. Therefore, the use of high-calorie, high-fat meals is recommended during food-effect bioavailability and fed bioequivalence studies.[64]

12.8.2 FDA guidance on food-effect bioavailability and bioequivalence studies

Food can change the bioavailability of a drug and can influence the bioequivalence between test and reference products. The influence of concomitant food intake on a drug prompted the FDA to issue a guidance for the industry on food-effect bioavailability and fed bioequivalence studies.[64]

In practice, it is difficult to determine the exact mechanism by which food changes the bioavailability of a drug product without performing specific mechanistic studies.[64] In terms of BCS class I drugs (high solubility/high permeability), important food effects on bioavailability are least likely to occur with many rapidly dissolving, IR drug products because absorption of the BCS class I drugs is usually pH- and site-independent and thus insensitive to differences in dissolution. However, for BCS class I drugs, food can still influence the bioavailability when there is a high first-pass effect, extensive adsorption, complexation, or instability of the drug substance in the GI; or, in some cases, excipients or interactions between excipients and food-induced changes in gut physiology can contribute to these food effects and influence the demonstration of bioequivalence; or, for rapidly dissolving formulations of BCS class I drug substances, food can affect drug peak concentration and the time to reach peak concentration by delaying gastric emptying and prolonging intestinal transit time. However, we expect the food effects on these measures to be similar for test and reference products in fed bioequivalence studies.

In terms of IR drug products for BCS class II, III, and IV drugs and for all modified-release drug products, food effects are most likely to result from a more complex combination of factors that influence the in vivo dissolution of the drug product, the absorption of the drug substance, or both. In these cases, the relative direction and magnitude of food effects on formulation bioavailability and the effects on the demonstration of bioequivalence are difficult, if not impossible, to predict without conducting a fed bioequivalence study.

Therefore, food effects have been investigated on BCS class II, III, and IV drugs. In general, positive food effects occur for BCS class II drugs when a higher systemic exposure is observed under fed conditions compared with the fasted state.[65,66] This is because food may improve absorption by enhancing the drug solubility in the intestinal lumen and inhibiting the efflux transporters in the intestine. Conversely, food can delay gastric emptying and prolong intestinal transit time, resulting in a delay in peak concentration of the drug product. In addition, in the presence of a meal, compounds of BCS class II weak bases are susceptible to variable dissolution and intestinal precipitation due to changes in GI pH and stimulated secretion of bile salts, while negative food effects are usually observed for most BCS class III compounds, showing decreased bioavailability after food intake.[66,67] For BCS class IV compounds, it is challenging to anticipate the direction of food effects because food effects can be formulation- and dose-dependent, as both solubility and permeability can be the rate-limiting steps for absorption.[66]

Based on the FDA guidance,[64] there are some recommendations for food-effect bioavailability and fed bioequivalence studies. For IR drug products, a food-effect bioavailability study should be conducted for all new chemical entities as part of IND applications and NDAs. Bioequivalence studies under fed conditions may be conducted as part of ANDAs for some IR drug products. Food-effect bioavailability and fed bioequivalence studies should be performed for all modified-release dosage forms.

12.9 REGIONAL DRUG ABSORPTION ALONG GI

It is well known that the transepithelial permeability of the small and large intestine varies with the intestinal site for drugs transported by passive diffusion, carrier-mediated mechanisms, or both. Understanding the regional differences in luminal conditions, GI transit and permeability is a fundamental requirement for establishing the subject-by-formulation interaction concept (which takes intersubject variations into consideration in relation to dosage form performance). Also, the permeability of each specific part of the intestine has to be taken into account in a more physiological manner in order to rationally design and develop advanced modified-release drug products. The information on regional drug absorption has been used to support and guide formulation development for drugs such as acipimox[68] and MK-0869.[69] Information of regional intestinal permeability in drug absorption along the GI have direct implications on the development of modified-release drug dosage forms.

12.9.1 Drug absorption from the stomach

Generally, absorption from the stomach is minor compared to from the intestine because of the gastric epithelial barrier preventing diffusion, the relatively small available surface area, and the negligible expression of nutrient (uptake) transport proteins. For a compound to be absorbed in the stomach, it needs to have high acidic solubility and rapid dissolution. Generally, the absorption of uncharged dissolved acidic drugs from the stomach is more ready than the absorption of basic drugs. Gastric acid secretion is one of the main parameters affecting the dissolution profile of solid dosage forms, particularly for drugs with pH-dependent solubility. The acidic secretion is controlled by various hormones and neurotransmitters, such as histamine, gastrin, acetylcholine, or somatostatin. Basal acid output amounts to 2–4 mmol/hour, but in a stimulated state, rates of up to 20–30 mmol/hour have been reported.[60] The gastric acid produced by the parietal cells has a maximum concentration of 160 mmol/L (pH 0.8). Under fasted conditions, the gastric pH value of healthy adults is reported to be within a pH range of 1–3. Under such acidic conditions, the dissolution of weak acids with higher pK_a values than the pH value of the gastric milieu (ie, ibuprofen, diclofenac, furosemide, and valproic acid) is hampered because they are primarily nonionized. In contrast, the dissolution of weak bases (ie, ketoconazole and dipyridamole) is enhanced because they are highly ionized. In addition, meal intake raises the pH values of the gastric content. Depending on meal composition and volume, values of up to pH 7 are reported in the literature, and the return to baseline pH values takes several hours, particularly for high-calorie meals.

12.9.2 Drug absorption from the small intestine

Small intestine is the primary site of absorption for most orally administered drugs. In addition, the proximal small intestine (P-SI) is more permeable than the more distal intestine (D-SI) and ascending colon. The effective permeation (P_{eff}) of drugs through the intestine often depends on the combined outcomes of passive diffusion and multiple parallel transport processes.[70] The human in vivo P_{eff} database from the jejunum is well established.[71] Single-pass perfusion studies have been used to calculate the P_{eff} of different substances from sites in the jejunum and ileum.[70] There is an in vivo human jejunal P_{eff} cutoff point at approximately 1.5×10^{-4} cm/second ($t_{1/2,abs}$ of about 42 minutes), which marks the difference between drugs that are rapidly absorbed (high P_{eff}) and those that are slowly and incompletely absorbed (low P_{eff}) from solution.[41] The cutoff point for the absorption rate based on in vivo jejunal perfusion data is in agreement with the simulated pharmacokinetic data.

Both intestinal metabolic enzymes and drug transporters determine the bioavailability of orally administrated drugs. One function of the small intestine in oral drug absorption can be attributed to its first-pass intestinal metabolism. The significant capability of the small intestinal mucosa to metabolize drugs has been reported, and the expression of metabolic enzymes (cytochrome P450s (1A1, 1B1, 2C9, 2C19, 2D6, 2E1, 2J2, 2S1, 3A4/5), UDP-glucuronosyltransferases, sulfotransferases, N-acetyltransferase) in human small intestine has been reported.[72] Intestinal metabolism mainly mediated by CYP can serve as a barrier to the systemic uptake of xenobiotics (including drugs) by facilitating excretion to the lumen of the intestine or by bioactivation of the xenobiotics, with consequent binding to enterocyte macromolecules. Another function of the small intestine in oral drug absorption can be attributed to drug

transport. The expression of a variety of influx (absorptive) transporters (intestinal oligopeptide transporter, dipeptide/tripeptide transporter (PepT1)) for amino acids, peptides, hexose, organic anions, organic cations, nucleosides, and other nutrients in human small intestine has been reported.[73,74]

In addition to influx transporters, adenosine triphosphate (ATP)−dependent efflux transporters such as MDR protein 1 (MDR 1 or P-glycoprotein 1), MDR-associated protein 2 (MRP2 or ABCC2), and breast cancer resistance protein (BCRP or ABCG2) are also expressed on the apical membrane of intestinal epithelia in humans and rodents.[73,74] P-gp is present at high levels in the villus enterocytes.[74] Also, the expression and activity of P-gp along the small intestine seems regional. In humans, the expression of P-gp is higher in the lower part of the intestine. P-gp has an affinity to a wide range of substrates to excrete various drugs into the lumen. It is an important factor that decreases the bioavailability of oral drugs, such as lumefantrine[75] or digoxin.[76] Notably, P-gp and CYP3A share a large number of substrates and inhibitors and are induced by many of the same compounds, including dexamethasone and rifampin.[77] Inhibitors of P-gp such as cyclosporine and verapamil have been shown to increase the drug absorptive transport.[78]

12.9.3 Drug absorption from colon

The colonic absorption of drugs can differ significantly from absorption in the small intestine as a consequence of several physiological, physicochemical, and biopharmaceutical factors. The permeability through passive transport is thought to be lower in colonic than in other intestinal tissue because of the smaller surface area and tighter junctions in the epithelial cell layer. There is a shortage of P_{eff} data from ileum and colon in the human intestinal tract.[71] There have not been any direct measurements of in vivo colonic P_{eff} for drugs in humans to date. In addition, the expression of efflux and uptake transporters such as P-gp and the human dipeptide/tripeptide transporter (hPepT1) appears to increase and decrease in the colon, which may limit the permeability, dissolution rate, and absorption. It has been reported that BCS class I drugs were generally well absorbed in the colon (with relative bioavailability in the colon greater than 70%) in a comparison of the plasma exposure of the drug following oral and colonic administration.[79] The low-P_{eff} drugs (BCS class III/IV) were less well absorbed in the colon (with relative bioavailability in the colon less than 50%). It has been reported that atenolol and metoprolol could function as P_{eff} markers for low- and high-colonic absorption, respectively.[79]

Interest has developed in directing drugs and dosage forms to effect primary drug release to the colon. Targeted distribution of an orally administered drug at the large intestine confers therapeutic advantages on the treatment of colonic diseases, peptide and protein therapy,[80] and chronotherapy.[81] To achieve such distribution control in the GI tract, the adoption of the prodrug concept gives birth to a colon-specific prodrug. Upon design of a colon-specific prodrug, one should take into consideration not only the delivery of a drug to the target site, but also its therapeutic effectiveness once it is there. The intestinal microbiota can be considered as a metabolically adaptable and rapidly renewable organ of the body. In addition, as knowledge will increase about the microbial bioconversion of polyphenolic compounds into bioactive metabolites in the colon, as well as whether food-based strategies can augment such bioconversion into more potent compounds with antioxidant and anti-inflammatory activity, new areas of research will be discovered. Furthermore, a colon-specific drug delivery system should prevent drug release in the stomach, as well as the small intestine. As the release of drugs from these polysaccharide-based systems is independent of pH and gastric-emptying time, these polysaccharide-based systems are considered the most effective and preferable means for colonic drug delivery in terms of target specificity.[82]

12.9.4 Advance in estimation of human in vivo regional intestinal permeability

The most recent study from the OrBiTo project (http://www.orbitoproject.eu/) presented the results for the indirect estimation of site-specific in vivo intestinal effective permeability (P_{eff}) in humans, and for the first time, it directly compared regional permeability along the human GI tract.[83] A good correlation was established between these indirect jejunal P_{eff} estimates and jejunal P_{eff} measurements determined directly using the single-pass perfusion double balloon technique. On average, P_{eff} estimates from the distal small intestine and large intestine were 90% and 40% of those from the proximal small intestine, respectively.

In this study, using deconvolution, the site specific P_{eff} can be calculated by

$$P_{eff} = \frac{CL_{abs}}{\text{absorptive surface area}} \tag{12.14}$$

where CL_{abs} is the intestinal absorption clearance rate.

Alternatively, P_{eff} can be expressed in a more straightforward way, assuming instantaneous and homogeneous distribution of active pharmaceutical

ingredients (APIs) in the total volume of the dose and subsequent flushes at all times:

$$P_{\text{eff}} = \frac{\text{absorption rate} \times \text{radius}}{\text{API remaining in lumen} \times 2} \quad (12.15)$$

12.10 FUTURE TRENDS

Significant advances have been made in the prediction of intestinal permeability using in silico techniques over the last decade. However, what is most needed now are more and better data to form the foundation of computational models that are capable of integrating the existing knowledge of intestinal permeability and permitting us to see relationships only hinted at in existing experiments. More and better data should also permit the application of nonlinear models to fit the observed curvilinear relationships between physicochemical properties/descriptors and permeability more accurately.

On the other hand, the effective use of cell culture techniques has dramatically improved the efficiency of permeability screening in the industry. Most biological screening assays are performed in miniaturized models such as 384- or 1536-well plates for ultrahigh-throughput screening.[84] Automation of cell-based drug absorption assays, including Caco-2 and MDCK, in 96-well format using permeable support systems has been reported.[85] Robotic systems coupled with the miniaturization devices can be extremely valuable because of the precision and high-speed handling of repetitive tasks and will definitely lead the way to an era of high-tech laboratories that operate smoothly, with minimal human intervention. Furthermore, some of the technological innovations can be made to reduce the time for cell growth. An automated 96-well plate Caco-2 permeation model with reduced growth time has been developed.[86] The efforts of improvements in culturing media, adjustments of seeding density, and use of modified transwell plates have already been made to shorten the cell growth time.[86] Further improvement in this area, with respect to cell culture techniques, will certainly be the focus of development efforts in the near future.

12.11 CONCLUSIONS

Optimizing bioavailability of orally administrated drug for systemic actions is one of the most important aims for the pharmaceutical industry. In order to achieve this goal, however, high-throughput, cost-effective, highly discriminating, and predictive models for the absorption of drug molecules in humans will be required. The current in silico approaches (physicochemical and certain in silico methods) seem to hold the promise of fulfilling this need, particularly as primary screening, with larger and more chemically diverse data sets ready before these models will be reliable enough to accurately predict the intestinal absorption.

Secondary screening can be achieved using automated, high-throughput in vitro systems (cell culture–based or animal tissue–based) for selecting and optimizing the chemical leads to identify potential drug candidates.

Finally, traditional approaches for the evaluation of drug absorption, such as in situ rat intestinal perfusion technique and in vivo animal or human studies, as well as PBPK models, should also be utilized simultaneously to prevent false positives and false negatives in the later stages of development.

DISCLAIMER

The opinions expressed in this report by the authors do not necessarily reflect the views or policies of the Food and Drug Administration (FDA).

References

1. *2015 Digestive System: Facts, Function & Diseases*. <http://www.livescience.com/22367-digestive-system.html>.
2. Helander HF, Fandriks L. Surface area of the digestive tract—revisited. *Scand J Gastroenterol* 2014;**49**:681–9.
3. Stephen AM, Cummings JH. The microbial contribution to human faecal mass. *J Med Microbiol* 1980;**13**:45–56.
4. Sellers RS, Morton D. The colon: from banal to brilliant. *Toxicol Pathol* 2014;**42**:67–81.
5. Takahashi T. Mechanism of interdigestive migrating motor complex. *J Neurogastroenterol Motil* 2012;**18**:246–57.
6. Bergstrom CA, Holm R, Jorgensen SA, Andersson SB, Artursson P, Beato S, et al. Early pharmaceutical profiling to predict oral drug absorption: current status and unmet needs. *Eur J Pharm Sci* 2014;**57**:173–99.
7. Mudie DM, Amidon GL, Amidon GE. Physiological parameters for oral delivery and in vitro testing. *Mol Pharm* 2010;**7**:1388–405.
8. Schiller C, Frohlich CP, Giessmann T, Siegmund W, Monnikes H, Hosten N, et al. Intestinal fluid volumes and transit of dosage forms as assessed by magnetic resonance imaging. *Aliment Pharmacol Ther* 2005;**22**:971–9.
9. Mudie DM, Murray K, Hoad CL, Pritchard SE, Garnett MC, Amidon GL, et al. Quantification of gastrointestinal liquid volumes and distribution following a 240 mL dose of water in the fasted state. *Mol Pharm* 2014;**11**:3039–47.
10. Benet LZ. The role of BCS (biopharmaceutics classification system) and BDDCS (biopharmaceutics drug disposition classification system) in drug development. *J Pharm Sci* 2013;**102**:34–42.
11. Torrado G, Carrascosa C, Torrado-Santiago S. Correlation of in vitro and in vivo acetaminophen availability from albumin microaggregates oral modified release formulations. *Int J Pharm* 2001;**217**:193–9.
12. Bredael GM, Bowers N, Boulineau F, Hahn D. In vitro-in vivo correlation strategy applied to an immediate-release solid oral dosage form with a biopharmaceutical classification system IV compound case study. *J Pharm Sci* 2014;**103**:2125–30.

13. FDA. *Guidance for industry: immediate release solid oral dosage forms scale-up and postapproval changes: chemistry, manufacturing, and controls, in vitro dissolution testing, and in vivo bioequivalence documentation.* <http://www.fda.gov/downloads/Drugs/Guidances/UCM070636.pdf>; 1995.
14. FDA. *Guidance for industry: dissolution testing of immediate release solid oral dosage forms.* <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm070237.pdf>; 1997.
15. FDA. *Guidance for industry: waiver of in vivo bioavailability and bioequivalence studies for immediate-release solid oral dosage forms based on a biopharmaceutics classification system.* <http://www.fda.gov/downloads/Drugs/Guidances/ucm070246.pdf>; 2015.
16. Emami J. In vitro—in vivo correlation: from theory to applications. *J Pharm Pharm Sci* 2006;**9**:169–89.
17. Awortwe C, Fasinu PS, Rosenkranz B. Application of Caco-2 cell line in herb-drug interaction studies: current approaches and challenges. *J Pharm Pharm Sci* 2014;**17**:1–19.
18. Luo Z, Morgan MR, Day AJ. Transport of trans-tiliroside (kaempferol-3-beta-D-(6″-p-coumaroyl-glucopyranoside) and related flavonoids across Caco-2 cells, as a model of absorption and metabolism in the small intestine. *Xenobiotica* 2015;**45**:1–9. Available from: http://dx.doi.org/10.3109/00498254.2015.1007492.
19. Kamei N, Onuki Y, Takayama K, Takeda-Morishita M. Mechanistic study of the uptake/permeation of cell-penetrating peptides across a caco-2 monolayer and their stimulatory effect on epithelial insulin transport. *J Pharm Sci* 2013;**102**:3998–4008.
20. Heinen C, Reuss S, Saaler-Reinhardt S, Langguth P. Mechanistic basis for unexpected bioavailability enhancement of polyelectrolyte complexes incorporating BCS class III drugs and carrageenans. *Eur J Pharm Biopharm* 2013;**85**:26–33.
21. Fernandez DA, Louzao MC, Vilarino N, Espina B, Fraga M, Vieytes MR, et al. The kinetic, mechanistic and cytomorphological effects of palytoxin in human intestinal cells (Caco-2) explain its lower-than-parenteral oral toxicity. *FEBS J* 2013;**280**:3906–19.
22. Larregieu CA, Benet LZ. Drug discovery and regulatory considerations for improving in silico and in vitro predictions that use Caco-2 as a surrogate for human intestinal permeability measurements. *AAPS J* 2013;**15**:483–97.
23. Fujikawa M, Ano R, Nakao K, Shimizu R, Akamatsu M. Relationships between structure and high-throughput screening permeability of diverse drugs with artificial membranes: application to prediction of Caco-2 cell permeability. *Bioorg Med Chem* 2005;**13**:4721–32.
24. Reis JM, Dezani AB, Pereira TM, Avdeef A, Serra CH. Lamivudine permeability study: a comparison between PAMPA, ex vivo and in situ Single-Pass Intestinal Perfusion (SPIP) in rat jejunum. *Eur J Pharm Sci* 2013;**48**:781–9.
25. Wang M, Zhang Y, Sun B, Sun Y, Gong X, Wu Y, et al. Permeability of exendin-4-loaded chitosan nanoparticles across MDCK cell monolayers and rat small intestine. *Biol Pharm Bull* 2014;**37**:740–7.
26. Zhang W, Parniak MA, Sarafianos SG, Empey PE, Rohan LC. In vitro transport characteristics of EFdA, a novel nucleoside reverse transcriptase inhibitor using Caco-2 and MDCKII cell monolayers. *Eur J Pharmacol* 2014;**732**:86–95.
27. Irvine JD, Takahashi L, Lockhart K, Cheong J, Tolan JW, Selick HE, et al. MDCK (Madin-Darby canine kidney) cells: a tool for membrane permeability screening. *J Pharm Sci* 1999;**88**:28–33.
28. Singh RS, Michel D, Das U, Dimmock JR, Alcorn J. Cytotoxic 1,5-diaryl-3-oxo-1,5-pentadienes: an assessment and comparison of membrane permeability using Caco-2 and MDCK monolayers. *Bioorg Med Chem Lett* 2014;**24**:5199–202.
29. Gres MC, Julian B, Bourrie M, Meunier V, Roques C, Berger M, et al. Correlation between oral drug absorption in humans, and apparent drug permeability in TC-7 cells, a human epithelial intestinal cell line: comparison with the parental Caco-2 cell line. *Pharm Res* 1998;**15**:726–33.
30. Li H, Chung SJ, Shim CK. Characterization of the transport of uracil across Caco-2 and LLC-PK1 cell monolayers. *Pharm Res* 2002;**19**:1495–501.
31. Adachi Y, Suzuki H, Sugiyama Y. Quantitative evaluation of the function of small intestinal P-glycoprotein: comparative studies between in situ and in vitro. *Pharm Res* 2003;**20**:1163–9.
32. Tavelin S, Taipalensuu J, Hallbook F, Vellonen KS, Moore V, Artursson P. An improved cell culture model based on 2/4/A1 cell monolayers for studies of intestinal drug transport: characterization of transport routes. *Pharm Res* 2003;**20**:373–81.
33. Tavelin S, Taipalensuu J, Soderberg L, Morrison R, Chong S, Artursson P. Prediction of the oral absorption of low-permeability drugs using small intestine-like 2/4/A1 cell monolayers. *Pharm Res* 2003;**20**:397–405.
34. Dvivedi J, Pandey S, Gupta R. Effect of curcumin on glucose absorption: an experimental study on albino rats. *Indian J Physiol Pharmacol* 2011;**55**:207–12.
35. Alam MA, Al-Jenoobi FI, Al-Mohizea AM. Everted gut sac model as a tool in pharmaceutical research: limitations and applications. *J Pharm Pharmacol* 2012;**64**:326–36.
36. Lennernas H. Animal data: the contributions of the Ussing Chamber and perfusion systems to predicting human oral drug delivery in vivo. *Adv Drug Deliv Rev* 2007;**59**:1103–20.
37. Li H, Jin HE, Shim WS, Shim CK. An improved prediction of the human in vivo intestinal permeability and BCS class of drugs using the in vitro permeability ratio obtained for rat intestine using an Ussing chamber system. *Drug Dev Ind Pharm* 2013;**39**:1515–22.
38. Fortuna A, Alves G, Falcao A, Soares-da-Silva P. Evaluation of the permeability and P-glycoprotein efflux of carbamazepine and several derivatives across mouse small intestine by the Ussing chamber technique. *Epilepsia* 2012;**53**:529–38.
39. Li H, Dong L, Liu Y, Wang G, Wang G, Qiao Y. Biopharmaceutics classification of puerarin and comparison of perfusion approaches in rats. *Int J Pharm* 2014;**466**:133–8.
40. Incecayir T, Tsume Y, Amidon GL. Comparison of the permeability of metoprolol and labetalol in rat, mouse, and Caco-2 cells: use as a reference standard for BCS classification. *Mol Pharm* 2013;**10**:958–66.
41. Lennernas H. Human in vivo regional intestinal permeability: importance for pharmaceutical drug development. *Mol Pharm* 2014;**11**:12–23.
42. Lennernas H, Fagerholm U, Raab Y, Gerdin B, Hallgren R. Regional rectal perfusion: a new in vivo approach to study rectal drug absorption in man. *Pharm Res* 1995;**12**:426–32.
43. Lipinski CA. Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov Today: Technol* 2004;**1**:337–41.
44. Sun L, Liu X, Xiang R, Wu C, Wang Y, Sun Y, et al. Structure-based prediction of human intestinal membrane permeability for rapid in silico BCS classification. *Biopharm Drug Dispos* 2013;**34**:321–35.
45. Nakao K, Fujikawa M, Shimizu R, Akamatsu M. QSAR application for the prediction of compound permeability with in silico descriptors in practical use. *J Comput Aided Mol Des* 2009;**23**:309–19.
46. Verma RP, Hansch C, Selassie CD. Comparative QSAR studies on PAMPA/modified PAMPA for high throughput profiling of drug absorption potential with respect to Caco-2 cells and human intestinal absorption. *J Comput Aided Mol Des* 2007;**21**:3–22.
47. Moda TL, Andricopulo AD. Consensus hologram QSAR modeling for the prediction of human intestinal absorption. *Bioorg Med Chem Lett* 2012;**22**:2889–93.
48. Iyer M, Tseng YJ, Senese CL, Liu J, Hopfinger AJ. Prediction and mechanistic interpretation of human oral drug absorption using MI-QSAR analysis. *Mol Pharm* 2007;**4**:218–31.
49. Suenderhauf C, Hammann F, Maunz A, Helma C, Huwyler J. Combinatorial QSAR modeling of human intestinal absorption. *Mol Pharm* 2011;**8**:213–24.

50. Talevi A, Goodarzi M, Ortiz EV, Duchowicz PR, Bellera CL, Pesce G, et al. Prediction of drug intestinal absorption by new linear and non-linear QSPR. *Eur J Med Chem* 2011;**46**:218–28.
51. Subramanian G, Kitchen DB. Computational approaches for modeling human intestinal absorption and permeability. *J Mol Model* 2006;**12**:577–89.
52. Kostewicz ES, Aarons L, Bergstrand M, Bolger MB, Galetin A, Hatley O, et al. PBPK models for the prediction of in vivo performance of oral dosage forms. *Eur J Pharm Sci* 2014;**57**:300–21.
53. EMA. 2014 *Concept paper on qualification and reporting of physiologically-based pharmacokinetic (PBPK) modelling and analyses*. United Kingdom. <http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/06/WC500169452.pdf>.
54. Huang W, Lee SL, Yu LX. Mechanistic approaches to predicting oral drug absorption. *AAPS J* 2009;**11**:217–24.
55. Yu LX, Amidon GL. A compartmental absorption and transit model for estimating oral drug absorption. *Int J Pharm* 1999;**186**:119–25.
56. Stoupi S, Williamson G, Viton F, Barron D, King LJ, Brown JE, et al. In vivo bioavailability, absorption, excretion, and pharmacokinetics of [14C]procyanidin B2 in male rats. *Drug Metab Dispos* 2010;**38**:287–91.
57. Takusagawa S, van Lier JJ, Suzuki K, Nagata M, Meijer J, Krauwinkel W, et al. Absorption, metabolism and excretion of [(14)C]mirabegron (YM178), a potent and selective beta(3)-adrenoceptor agonist, after oral administration to healthy male volunteers. *Drug Metab Dispos* 2012;**40**:815–24.
58. Abel S, Russell D, Whitlock LA, Ridgway CE, Nedderman AN, Walker DK. Assessment of the absorption, metabolism and absolute bioavailability of maraviroc in healthy male subjects. *Br J Clin Pharmacol* 2008;**65**(Suppl. 1):60–7.
59. Mahmood I. Estimation of absorption rate constant (ka) following oral administration by Wagner-Nelson, Loo-Riegelman, and statistical moments in the presence of a secondary peak. *Drug Metabol Drug Interact* 2004;**20**:85–100.
60. Koziolek M, Garbacz G, Neumann M, Weitschies W. Simulating the postprandial stomach: physiological considerations for dissolution and release testing. *Mol Pharm* 2013;**10**:1610–22.
61. Custodio JM, Wu CY, Benet LZ. Predicting drug disposition, absorption/elimination/transporter interplay and the role of food on drug absorption. *Adv Drug Deliv Rev* 2008;**60**:717–33.
62. Dahan A, Altman H. Food-drug interaction: grapefruit juice augments drug bioavailability--mechanism, extent and relevance. *Eur J Clin Nutr* 2004;**58**:1–9.
63. Arayne MS, Sultana N, Bibi Z. Grape fruit juice-drug interactions. *Pak J Pharm Sci* 2005;**18**:45–57.
64. FDA. 2002 *Guidance for industry: food-effect bioavailability and fed bioequivalence studies*. FDA. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM070241.pdf>.
65. Xia B, Heimbach T, Lin TH, Li S, Zhang H, Sheng J, et al. Utility of physiologically based modeling and preclinical in vitro/in vivo data to mitigate positive food effect in a BCS class 2 compound. *AAPS PharmSciTech* 2013;**14**:1255–66.
66. Heimbach T, Xia B, Lin TH, He H. Case studies for practical food effect assessments across BCS/BDDCS class compounds using in silico, in vitro, and preclinical in vivo data. *AAPS J* 2013;**15**:143–58.
67. Heinen CA, Reuss S, Amidon GL, Langguth P. Ion pairing with bile salts modulates intestinal permeability and contributes to food-drug interaction of BCS class III compound trospium chloride. *Mol Pharm* 2013;**10**:3989–96.
68. Menon R, Cefali E, Wilding I, Wray H, Connor A. The assessment of human regional drug absorption of free acid and sodium salt forms of acipimox, in healthy volunteers, to direct modified release formulation strategy. *Biopharma Drug Dispos* 2009;**30**:508–16.
69. Wu Y, Loper A, Landis E, Hettrick L, Novak L, Lynn K, et al. The role of biopharmaceutics in the development of a clinical nanoparticle formulation of MK-0869: a Beagle dog model predicts improved bioavailability and diminished food effect on absorption in human. *Int J Pharm* 2004;**285**:135–46.
70. Lennernas H. Regional intestinal drug permeation: biopharmaceutics and drug development. *Eur J Pharm Sci* 2014;**57**:333–41.
71. Dahlgren D, Roos C, Sjogren E, Lennernas H. Direct in vivo human intestinal permeability (P) determined with different clinical perfusion and intubation methods. *J Pharm Sci* 2014;**104**:2702–26. Available from: http://dx.doi.org/10.1002/jps.24258.
72. Zhao Y, Hebert MF, Venkataramanan R. Basic obstetric pharmacology. *Semin Perinatol* 2014;**38**:475–86.
73. Takano M, Yumoto R, Murakami T. Expression and function of efflux drug transporters in the intestine. *Pharmacol Ther* 2006;**109**:137–61.
74. UCSF-FDA. *UCSF-FDA Transportal*. <http://dbts.ucsf.edu/fdatransportal/organs/small-intestine/>.
75. Wahajuddin, Raju KS, Singh SP, Taneja I. Investigation of the functional role of P-glycoprotein in limiting the oral bioavailability of lumefantrine. *Antimicrob Agents Chemother* 2014;**58**:489–94.
76. Igel S, Drescher S, Murdter T, Hofmann U, Heinkele G, Tegude H, et al. Increased absorption of digoxin from the human jejunum due to inhibition of intestinal transporter-mediated efflux. *Clin Pharmacokinet* 2007;**46**:777–85.
77. Roberts PJ, Rollins KD, Kashuba AD, Paine MF, Nelsen AC, Williams EE, et al. The influence of CYP3A5 genotype on dexamethasone induction of CYP3A activity in African Americans. *Drug Metab Dispos* 2008;**36**:1465–9.
78. Liu ZQ, Jiang ZH, Liu L, Hu M. Mechanisms responsible for poor oral bioavailability of paeoniflorin: Role of intestinal disposition and interactions with sinomenine. *Pharm Res* 2006;**23**:2768–80.
79. Tannergren C, Bergendal A, Lennernas H, Abrahamsson B. Toward an increased understanding of the barriers to colonic drug absorption in humans: implications for early controlled release candidate assessment. *Mol Pharm* 2009;**6**:60–73.
80. Pawar VK, Meher JG, Singh Y, Chaurasia M, Surendar Reddy B, Chourasia MK. Targeting of gastrointestinal tract for amended delivery of protein/peptide therapeutics: strategies and industrial perspectives. *J Control Release* 2014;**196**:168–83.
81. Lee Y, Kim IH, Kim J, Yoon JH, Shin YH, Jung Y, et al. Evaluation of dextran-flufenamic acid ester as a polymeric colon-specific prodrug of flufenamic acid, an anti-inflammatory drug, for chronotherapy. *J Drug Target* 2011;**19**:336–43.
82. Pachuau L, Mazumder B. Colonic drug delivery systems based on natural polysaccharides and their evaluation. *Mini Rev Med Chem* 2013;**13**:1982–91.
83. Sjogren E, Dahlgren D, Roos C, Lennernas H. Human in vivo regional intestinal permeability: quantitation using site-specific drug absorption data. *Mol Pharm* 2015;**12**:2026–39.
84. Miyamoto R, Nozawa T, Kimura M, Shiozuka K, Tabata K. Development and validation of semiautomated 96-well transport assay using LLC-PK1 cells transfected with human P-glycoprotein for high-throughput screening. *Assay Drug Dev Technol* 2015;**13**:79–87.
85. Larson B, Banks P, Sherman H, Rothenberg M. Automation of cell-based drug absorption assays in 96-well format using permeable support systems. *J Lab Autom* 2012;**17**:222–32.
86. Galkin A, Pakkanen J, Vuorela P. Development of an automated 7-day 96-well Caco-2 cell culture model. *Pharmazie* 2008;**63**:464–9.

CHAPTER

13

Dissolution Testing of Solid Products

Y. Chen[1], *Z. Gao*[2] *and J.Z. Duan*[2]

[1]Novast Laboratories, Ltd., Nantong, China [2]U.S. Food and Drug Administration, Silver Spring, MD, United States

13.1 INTRODUCTION

A dissolution test is a comparative tool for measuring the in vitro performance of solid oral dosage forms. Dissolution of a solid product can be affected by the compositions and the properties of the powder, including particle size, polymorphism of the active pharmaceutical ingredients (APIs), excipients in the formulated product, the manufacturing process, and factors of the method and environment for dissolution testing. With proper selection of the test method in a well-controlled environment, a dissolution test can be sensitive to changes in composition, material property, and the manufacturing process. Therefore, dissolution testing is commonly used to design and select formulations and as a comparative tool for quality control for batch-to-batch consistency in routine manufacturing of solid pharmaceutical products.

When the human pharmacokinetic (PK) data are available from a sufficient number of versions of a product with different dissolution profiles, a properly developed dissolution test may lead to the development of in vitro–in vivo correlation (IVIVC)[1–3] for the drug products. In these cases, the dissolution test may be used as a surrogate for the in vivo performance test. With a clear correlation of in vitro dissolution with in vivo performance, a dissolution test can be used for a waiver of bioequivalence (BE) study in humans for registration of certain type of rapidly releasing biopharmaceutics classification system (BCS) Class 1 or 3 products,[4] and for significant changes in scale-up and postapproval changes (SUPACs)[5,6] via the use of the established IVIVC. Thus, such a dissolution test could decrease the regulatory agency review time and the pharmaceutical industry cost by reducing the development time and costs for the in vivo BE studies.

Dissolution testing has been applied as compendial test for solid oral products in the United States Pharmacopoeia (USP) since the late 1960s. Multiple chapters related to dissolution testing, such as Chapters <711>, <724>, <1088>, <1090>, and <1092>, are included in the current version of USP. In addition, dissolution assays have been adopted by regulatory agencies in different countries[2,7–11] as one of the quality requirements for approval of new products of solid oral dosage forms either by new drug applications (NDAs) or by abbreviated new drug applications (ANDAs) for generic drugs. Currently, all solid oral dosage forms are required to have a dissolution test, except for some cases in which disintegration may be more appropriate than dissolution as a performance test.

A dissolution test can be used to guide product development and for quality control of products. When an effective IVIVC can be established, the in vitro dissolution test is a valid measure for the in vivo quality/performance of the product. By contrast, a dissolution test is not a valid measure of in vivo performance when IVIVC has not been established. Furthermore, the same dissolution results for multisource products do not guarantee that the products will have the same in vivo performance. In the absence of IVIVC, the use of dissolution data is ambiguous for comparing different products.

Unlike other product quality attributes, which can be reproduced by different test methods, product dissolution is highly method dependent. Significantly different dissolution data can be generated for the same product using different dissolution methods, with all of the different results being true for the same product. Without IVIVC, the decision on the selection of test procedures and setting of dissolution specification

DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00013-3

becomes the preference of the people involved. In addition, there are different views and interpretations of dissolution test data by the regulatory agencies from different countries. Thus, the requirement of a dissolution test for drug products varies from country to country.

As indicated previously, dissolution data from a drug product can be affected by many factors of the dissolution test method, including the compositions of the test medium, test apparatus, mixing speed, aeration of the medium, and the physical environment for the test. Therefore, the contributions and influence of these factors must be carefully measured to achieve the proper interpretation and application of dissolution tests.

This chapter will begin with a brief review of the theory of dissolution, followed by an in-depth discussion on the current technologies for dissolution testing and suggestions for regulatory requirements of the test for products. The development and application of IVIVC will be discussed in chapter "In Vitro/In Vivo Correlations: Fundamentals, Development Considerations, and Applications," so it will not be discussed here.

Compendial requirements and Food and Drug Administration (FDA) guidance are referenced extensively, but not comprehensively, in this chapter. The authors acknowledge that different countries have different requirements. In fact, readers are encouraged to compare the differences between countries to ensure that the appropriate regulatory expectations are met based on the intended use and markets, and to promote the science and the proper use of dissolution tests.

13.2 THEORY OF DISSOLUTION TEST FOR SOLID DRUG PRODUCTS

13.2.1 Dissolution and drug absorption

After oral administration of any solid oral dosage forms (eg, immediate release (IR) or modified release (MR) tablets or capsules), the APIs in the dosage forms must dissolve into the gastrointestinal (GI) fluid (solution) before the APIs can become therapeutically effective (eg, locally in the GI tract or after being absorbed into the bloodstream to reach the site of action), as shown in Fig. 13.1.

The dissolution rate from the dosage form into the GI fluid may affect the drug absorption rate, hence the therapeutic effect of the dosage forms. Dokoumetzidis and Macheras[12] in their review of the history of dissolution testing credited Edwards[13] with publishing the first studies describing the relationship

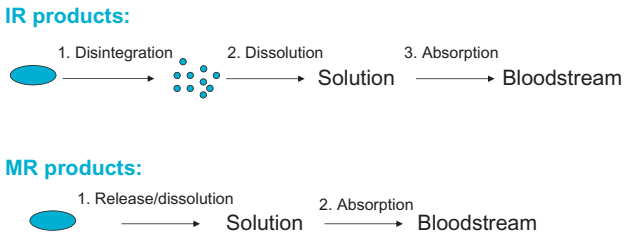


FIGURE 13.1 Schematic of the absorption steps of solid oral dosage forms.

between rates of dissolution and absorption. Initially, dissolution testing was included in the USP with an intention of measuring the physiological availability of the product in the GI fluid. However, direct verification of the dissolution in the GI fluid is difficult. Notably, the GI fluid is not the final destination for drugs that are entering the systemic circulation. After a drug dissolves into the physiological fluid in the GI tract, the drug will pass through the GI membrane and be transported into the bloodstream to reach the site of action and exert a pharmacological effect. Assuming that the absorption mechanism is the same for different lots of the same product or the same product made by different manufacturers (generics), the dissolution in the GI tract can be inferred by measuring the bioavailability (BA) of the product after oral administration. *Bioavailability* is defined as the rate and extent to which the drug is absorbed from a drug product and becomes available at the site of action [21 CFR 320.1(a)]. The concept of relating in vitro dissolution to in vivo BA was officially incorporated in USP19.

Over the past 50 years, the interest in developing in vitro tools to describe and attempt to predict in vivo performance, or the bioperformance, has been a central focus of biopharmaceutical research. In a number of studies, Amidon and his coworkers have developed comprehensive models to describe the transit of drug suspensions through the GI tract using chemical engineering concepts coupling mass transfer and fluid dynamics.[14,15] These models established the basic relationships between drug solubility, dissolution rate, and GI membrane permeability. New dimensionless parameters were introduced, including the dose number (which relates the dose applied to the ability to solubilize the ingested dose) and the absorption number (which relates the drug permeability to the residence time). These models were the basis for establishment of the BCS[16] that has provided a framework for interpretation of dissolution studies in the regulatory approval process, particularly in biowaivers for approval of new drug products and for justification of significant changes made to existing products via the

established IVIVC.[4–6] Over the last half-century, numerous methods have been tested in the quest to establish predictive methods for drug absorption from solid oral products. While there has been some successful examples, IVIVC is still not yet been established for many products due to the complexity of the in vivo dissolution and absorption processes. The dissolution process in the GI tract is not yet fully understood. An established IVIVC for one product is not necessarily applicable to a second product, not even for multisource products. Therefore, more study is required to understand the differences between the in vitro and in vivo dissolution processes. Once the differences are defined, the in vitro dissolution methods may predict the bioperformance of solid drug products, and to increase the regulatory flexibility for certain formulation and process changes.[4–6]

13.2.2 Dissolution tests for quality control

From the history of dissolution testing for oral solid products, it is known that product dissolution profiles can be affected by the quantitative and qualitative drug composition differences, manufacturing processes, the physical properties of the active and inactive ingredients and the various parameters of dissolution test methods. Thus, a change in compositions may lead to a change in product dissolution profiles. The responses of product dissolution to other changes in manufacturing of a product with fixed quantitative compositions increase the possibility that a dissolution test could become a useful tool for product quality control if the dissolution method is established to have sufficient robustness for monitoring the changes. For example, the change in the powder properties of a given formulation, such as the particle size distributions of API/excipients/granules,[17–19] polymorphism of API,[17,20,21] density of granules, viscosity grade of polymers, and source of excipients,[22] could lead to significant changes in dissolution profiles. The difference in manufacturing processes, such as direct compression and wet granulation, granulation conditions/parameters, lubrication time,[23] compression force,[18] and coating process parameters,[24] may also lead to different dissolution profiles. Some of the changes in dissolution may lead to changes in their bioperformance of the products, while others may not. That is, IVIVC could not be established for all products with all the changes. However, the sensitivity and versatile capability of dissolution testing to detect the material and process changes can be utilized as a tool for quality control by monitor the consistency in material and process, which may be advantageous over other existing analytical techniques.

The following aspects should be considered when using a dissolution test as a tool for quality control in the absence of a valid IVIVC:

1. The value of dissolution data does not have the same clear-cut implications for quality as the results of other types of tests such as potency or purity, in that the dissolution data is method dependent, while the other tests are not. For example, different dissolution methods/conditions may result in significantly different dissolution data for a same product, but the assay of a product can be the same using different test methods.
2. Dissolution tests can be meaningful for postmarketing quality control by vertically comparing with the specifications set based on the dissolution data of the BE/BA batches of the same product, but not for lateral comparison with different products in most cases except for products that are qualified for biowaiver. Other quality attributes, such as assay and impurities, can be used to compare laterally for different products in addition to the vertical comparison for the history of the same product.
3. If the method is not developed properly, dissolution results can be sensitive to test conditions, as well as to the manufacturing process for drug products. Therefore, the changes in dissolution data may be due to the lack of robustness of the method itself rather than the true significant changes in the manufacturing or in the bioperformance of the product.
4. While it is desirable to have discriminating power to detect significant changes that could lead to bioinequivalence of a product, dissolution testing should be robust to tolerate the normal variability in routine manufacturing to prevent unnecessary rejection of products. The need of discriminating power for quality control and the robustness for daily operation should be balanced during method development for each product.

13.2.3 Mechanism of dissolution

More than a century ago, Noyes and Whitney developed a basic equation indicating that when surface area is constant, the dissolution rate of a solid is proportional to the difference between the solubility and concentration in the bulk solution.[25] Later, Nernst and Brunner proposed a diffusion layer theory to correlate dissolution rate with the diffusion coefficient, surface area, solubility, and the thickness of the diffusion layer.[26,27] The Noyes–Whitney and the Nernst–Brunner equations formed the fundamental diffusion layer theory for studying particle dissolution. There

are also diffusion-convective models using various mathematical equations for dissolution process under different assumptions. Details on these models are discussed in chapter "Fundamental of Diffusion and Dissolution."

Hypothetically, the dissolution process for a solid may include three steps: (1) detachment of drug molecules from the solid, (2) solvation of the detached drug molecules by solvent (medium) molecules in a stagnant layer on the surface of the solid, and (3) transporting the solvated drug molecules from the stagnant layer into the bulk solution. The rates of the hypothetical steps 1 and 2 are governed by the thermodynamic properties of the solid, including the enthalpy and entropy of melting, and that of molecular interaction between the solute and the solvent. The correlation between the thermodynamic properties with solubility (hence the dissolution rate) has been well studied.[28] For the transport process in step 3, the thermodynamic properties for molecular interaction continue to be the governing factor for the diffusion process. However, pure diffusion is a very slow process that is not suitable for routine testing. In modern dissolution tests when mixing is used, the transport mechanism becomes a mixed type of diffusion and convection, in which convection may be the dominant mechanism for the transport. In other words, mixing rate (hence the hydrodynamic of the medium) is the key factor responsible for stripping the solvated drug molecules from the solid surface and transporting the drug molecules into the bulk solution in dissolution test. Therefore, dissolution rate is affected by mixing speed. See chapter "Fundamental of Diffusion and Dissolution" for different mathematical expressions relating dissolution rates to the hydrodynamic parameters under varied assumptions and conditions.

13.3 CURRENT TECHNOLOGY AND INSTRUMENTATION FOR DISSOLUTION TESTING

13.3.1 Current USP dissolution apparatus for oral dosage forms

There are seven apparatuses defined in USP general Chapter <711>. This chapter was recently harmonized with the European Pharmacopeia (Ph.Eur.) and Japanese Pharmacopeia (JP) to create common definitions and operating parameters for all drug release monitoring. Of the seven apparatuses, Apparatus V (paddle over disk) and Apparatus VI (rotating cylinder) are not used for oral drug products and will not be discussed here.

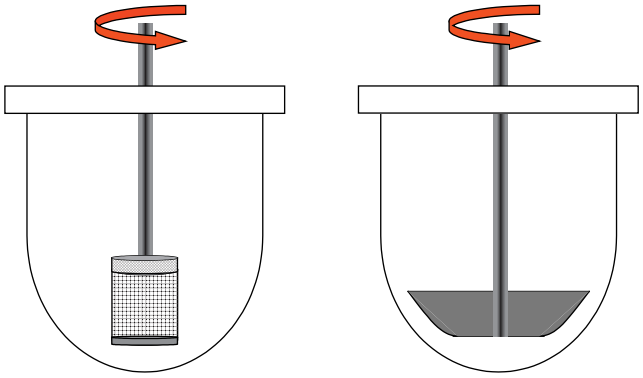


FIGURE 13.2 Schematic illustration of dissolution Apparatus I (basket) and II (paddle).

The widely used apparatuses for dissolution testing of solid oral dosage forms are Apparatus I and Apparatus II (Fig. 13.2). Both apparatuses employ a single hemispherical bottom vessel containing the dissolution medium with a centered spindle. The temperature of the vessel is controlled within narrow limits (±0.5°C) by immersion in constant temperature water bath or by other suitable technology. For Apparatus I, the dosage form is placed in a basket affixed to the end of the spindle. Agitation is achieved by rotating the basket; a rotational speed of 100 rpm is most commonly used. For Apparatus II, the dosage form is dropped directly into the vessel, and a paddle is used to induce agitation with typical rotation speeds of 50–75 rpm. In the cases where the dosage form may float (eg, capsule formulations), Apparatus I may be more appropriate. Alternatively, when using Apparatus II, the dosage form, either tablet or capsule, may be placed in a sinker designed to enclose the dosage form. There are various sinker designs (Fig. 13.3), and the most common sinker, as described by USP Chapters <711> and <1092>, consists of "a few turns of a wire helix." Soltero et al.[29] demonstrated that sinker design could affect dissolution results by entrapping material or altering the hydrodynamics and slowing dissolution. In addition to inhibiting flotation, the sinker may be useful for assuring consistent placement of the dosage form in the center of the vessel bottom. The hydrodynamic environment in the Apparatus II vessel is highly heterogeneous,[30,31] and it has been demonstrated that the position of the dosage form (eg, side vs center of the vessel or height of the dosage form) can alter the rate of dissolution.[32,33] For both Apparatuses I and II, the typical medium volume are 500 and 900 mL. Smaller and larger vessels are available to accommodate either low-dose or poorly soluble compounds, respectively. The medium type and agitation speed are selected as part of the method development process to whatever dissolution medium

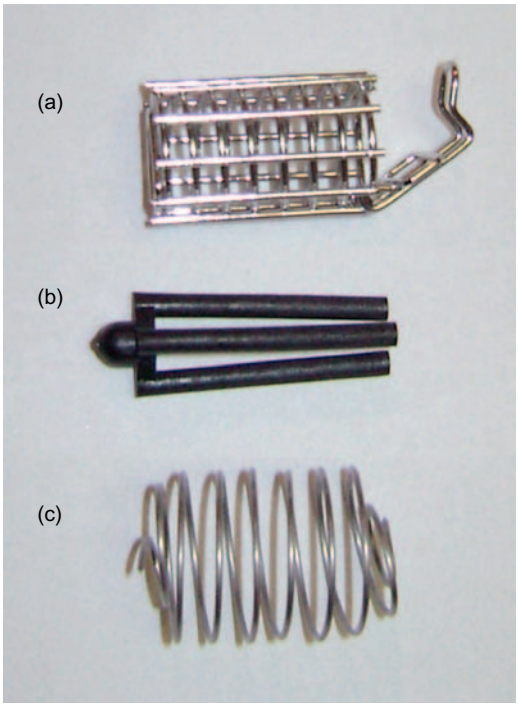


FIGURE 13.3 Example sinker designs. (a) Basket sinker, (b) three-pronged sinker, and (c) spiral sinker.

and agitation speed can be justified as providing an appropriate measure of product performance.

Sampling from Apparatuses I and II can be conducted either manually or automatically with filtration to eliminate undissolved drug particles and insoluble ingredients. The sampling time is recorded as the point of initiation of sampling. Manual sampling is conducted using syringes equipped with cannulas of a prescribed length to ensure sampling within the USP sampling zone, defined vertically as midway between the medium surface and the top of the paddle blade or basket, and radially as not less than 1 cm from the vessel wall. Dissolution vessels are typically fitted with covers containing sampling ports to facilitate reproducible sampling. These covers also minimize evaporation of the test medium during extended trials.[34] Automatic sampling devices employ sampling manifolds that either deliver the sample to a fraction collector or pass the aliquot through an online detector. Filtration can occur at the sampling port or inline. Whether manual or automatic methods are used, sampling can be conducted with or without replacement of the withdrawn media. With online detectors, the sampled aliquot can be returned to the vessel. In other cases, an equivalent volume of prewarmed media is delivered carefully to the vessel to prevent unacceptable alteration of the vessel hydrodynamics. Calculation of the amount of drug released versus time must be corrected for the volume changes that occur when the medium is not replaced.

Less commonly used equipment for measuring drug release from solid oral products includes Apparatuses III, IV, and VII. Apparatus III uses a reciprocating cylinder to induce agitation and to transport the dosage form through a series of vessels. This design allows monitoring release through a series of media at different pH to simulate the pH gradient of the GI tract. The reciprocating action may also be a better model of the pulsatile hydrodynamics of the GI tract than the hydrodynamic environments of Apparatus I or II.[35] Apparatus VII is a similar design, providing the capability of using smaller vessel volume (approximately 10 mL, as opposed to 250 mL for Apparatus III) and a shorter stroke length. Apparatus VII is used for extended release (ER) products where the amount of drug released is low and the smaller vessel volume improves the ability to quantify reliably the amount of drug released. Apparatus IV is a flow-through system designed to have improved control over the hydrodynamics of the medium. The flow has a pulsatile motion to mimic intestinal flow. This design also allows for continuous changes in the medium, such as inducing a pH gradient.

13.3.2 Possible variables during dissolution testing

Dissolution testing is widely used as an analytical technique for evaluating the drug release characteristics and consistency of a pharmaceutical product. A dissolution method for any particular product, including apparatuses, speeds, and media, is developed based on the characteristics of the product. However, dissolution results from the same product using the same method can still be affected by test variables. Thus, minimizing the variance due to the system used and the operator performing the work to obtain results that distinguish changes between different dosage forms or the same dosage form from different lots is necessary. Dissolution testing involves many variables, which can be grouped into four main categories: (1) analyst, (2) dissolution apparatus, (3) testing environment, and (4) sample.[36] The variables in the second category include vessel geometric irregularities, alignment of the instrument, and agitation speed. Variables of the testing environment include deaeration of dissolution media, media temperature, media pH, and vibration. To obtain reliable dissolution results, the variables from the first three categories must be minimized.

Gauge repeatability and reproducibility (gauge R&R) is a powerful tool that has been used to analyze

variability for USP Apparatus II dissolution measurement systems. The studies were designed to assess the variability due to apparatus, operator, and sample tablet. Evaluation of dissolution test results at 30 minute using 10-mg prednisone tablets (FDA/DPA NCDA#2) indicates that in the main contribution to the total variance, approximately 70% is due to the sample tablets, approximately 25% is from the apparatus, and approximately 5% is due to the operators. The results demonstrate that, following the American Society for Testing and Materials (ASTM)[37] procedure for mechanical calibration, the variation (gauge) due to a measurement system is small relative to the variation in the samples.

For some products, dissolved gases in the dissolution medium affect dissolution results. Thus, as the FDA Guidance for Industry states, these products require degassing of the medium prior to use for reproducible results. The solubility of nitrogen and oxygen, the main components of air, in water decreases with increasing temperature. This change in solubility causes a net flux of gas out of the water if the temperature is raised. If using a medium that has not been sufficiently degassed, bubbles can be seen forming around a tablet or capsule, which has been introduced to the dissolution vessel. For some products, these bubbles can affect dissolution results. A study conducted by the FDA suggested that measurement of total gases and not just oxygen in the medium was necessary to assess adequacy for dissolution testing.[38] The results indicate that the percent saturation of the total dissolved gases at room temperature must be well below 100% to prevent outgassing once the medium is brought to the dissolution test method temperature (typically 37°C). A general guidance for sufficient degassing is shown in Fig. 13.4. If the measured percent saturation of total dissolved gases at a given degassing temperature is below the solid line, outgassing will not occur when the temperature of the dissolution media is brought to 37°C.

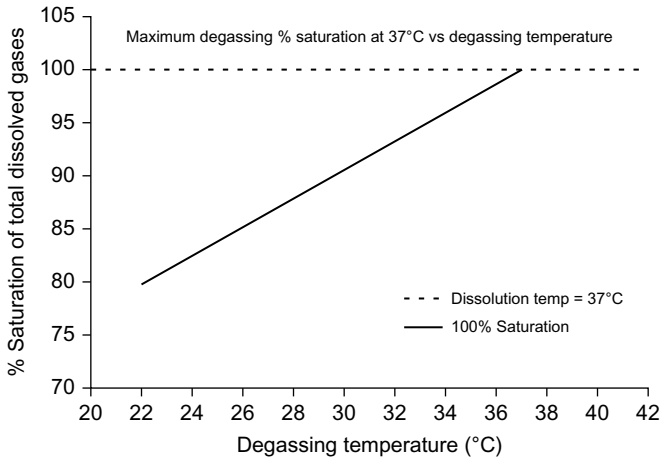


FIGURE 13.4 Maximum percent saturation of total dissolved gases allowed to avoid outgassing at 37°C.

The USP specifications in general Chapter <711> for a dissolution vessel currently specify the inner diameter of the cylindrical part and the height of vessel and assume that the inner surface of cylinder and hemisphere of the vessel are uniform.

The FDA conducted a study to explore the effect of the inner shape of standard 1-L dissolution vessels on drug dissolution results.[39] The geometric dimensions and irregularities of commercially available vessels (obtained from four different manufacturers) were examined using a three-dimensional video-based measuring machine. The definitions of geometric dimension and irregularity are included in Table 13.1.

The results showed that geometric characteristics varied within and among the sets of vessels, but the overall averages and standard deviations (SDs) of dissolution results (six vessels) showed no statistical significant differences among the vessel sets. The similarity of dissolution results indicates that when the dissolution apparatus is set up correctly and mechanically calibrated following the ASTM procedure, the variability can be minimized. However, some dissolution differences were noted when comparing individual vessels. With these types of comparisons, the vessel concentricity, sphericity, and radius of sphere were found to possibly influence the amount of drug dissolved, but flatness of vessel flange, cylindricity, and circularity showed no effect on dissolution results.

On the question regarding vibration effects on the dissolution experiment, the USP simply states that "no part of the (dissolution) assembly, including the environment in which the assembly is placed, contributes significant motion, agitation, or vibration." A vibration is a periodic motion, or one that repeats itself after a certain interval of time. Vibration may be generated by motors, mechanical motions, construction work, or vehicle traffic and can be transmitted through walls, floors, lab benches, and apparatuses. The vibration encountered during dissolution measurement includes system and environmental vibration. For a mechanically calibrated and well maintained dissolution apparatus, system vibration can be minimized. However, environmental vibration induced by laboratory equipment, building construction, or even by the analysts themselves is one of the more complicated factors affecting dissolution testing. The studies suggest that vibration affects the disintegrating and dissolving processes of drug products through different mechanisms. The results with 10-mg prednisone tablets indicate that applied vibration can lead to either an increase or a decrease in dissolution results, depending upon when that vibration occurs and the magnitude of the

TABLE 13.1 Definitions of Geometric Dimensions and Irregularities

Terms		Definitions
Geometric dimensions	Sphere radius	The measured value of the radius of the hemispheric region.
	Vessel height	The measured value from the base of the vessel to the top of the flange.
	Angle of cylinder to flange	The angle measured between the flange plane (horizontal) and the axis (vertical) of the cylinder.
	Diameter of cylinder	The measured value of the perpendicular cross-section diameter across the cylinder.
Geometric irregularities	Sphericity	All points on a three-dimensional (3D) surface share a common center point. The value will be zero for perfect shape.
	Circularity	The condition on a two-dimensional (2D) surface such as a sphere, cone, or cylinder where all the points along that surface are of equal distance to the center point of that feature. In the case of a circle, the points would be of equal distance to the center of the circle. For the perfect shape, the value will be zero.
	Cylindricity	All points on a surface are of equal distance from a common axis.
	Concentricity of two features (sphere part and cylinder part for dissolution vessel)	Both features share a common axis or center point. The measured sphere and cylinder share a common axis.
	The perpendicularity of the cylinder to the flange	The condition of the center axis of the cylinder as related to the 90-degree axis of the plane formed by the flange. For the perfect shape, the value will be zero.
	The flatness of the flange	All points measured are in the same plane. The plane is formed by all points measured on the flange surface. The flatness measurement is the sum of two distances from the plane. Distance 1 is the distance from the plane to the point farthest away from the plane in a positive direction. Distance 2 is the distance from the plane to the point farthest away from the plane in a negative direction. For the perfect shape, the value will be zero.

vibration during the dissolution test. These are particularly problematic because they may be beyond the analyst's control and typically cannot be predicted. The vibration measurement system described in the literature enables vibration changes occurring during dissolution testing to be monitored.[40–42]

Understanding how operational variables affect dissolution results will help drug development and product quality control by ensuring that the variables are properly controlled. For example, if a formulation is intended for IR, sample centering may need to be considered when using the USP paddle method. By contrast, if the formulation is an ER drug, in addition to all the other mechanical parameters, environmental vibration may need to be controlled during the dissolution testing.

Several variables, including dissolved gases in the dissolution medium, off-center placement of the test tablet, environmental vibration and various agitation speeds have been modeled. Mathematical models including Higuchi, Korsmeyer–Peppas, Weibull, and the Noyes–Whitney equation were employed to study the dissolution profile of 10-mg prednisone tablets (NCDA #2) using the USP paddle method. The results showed that dissolution variables affected dissolution rate constants differently, depending on whether the tablets disintegrated or dissolved. The detailed studies indicated that different dissolution variables affect drug release mechanisms differently. Placing a sample off-center caused a hydrodynamic change and mainly affected drug release during the disintegration process. Dissolved gases in the dissolution medium showed a complicated effect on the dissolution profiles, which may depend on the degree of gas saturation and the speed of bubble formation during dissolution testing, as well as interactions between the drug substance and the bubble surface. Environmental vibrations changed the dissolution rate by changing the aggregate condition of the disintegrated particles. Compared with the agitation in the medium, induced vibrations could produce a "secondary movement," which was relatively small and did not change the disintegration rate significantly, but did affect drug release rate by diffusion.

13.3.3 Calibration of dissolution apparatus

Currently, the qualification process for ensuring that a paddle or basket apparatus is suitable for its intended use is a highly debated and controversial

topic. Different instrument qualification and suitability methods have been proposed by various pharmacopeias and regulatory bodies.[43] The USP includes mechanical calibration specifications and the use of performance verification testing (PVT) to establish instrument suitability.[44] The process by which a test apparatus is determined to meet the compendial dimensional and operational specifications has been termed *mechanical calibration*, while the use of reference standard tablets is given the designation *chemical calibration*.[45] According to USP, the suitability of the dissolution paddle or basket assembly is determined by conformance to the dimensions and tolerances stated in its Chapter <711>. Chemical calibration allows a final and summative confirmation of the suitable operation of the integrated dissolution assembly, beyond the evaluation of its separate component attributes (eg, stirring element dimensions, control of rotation speed, dissolution medium volume). USP thinks that neither mechanical nor chemical calibration alone is sufficient to adequately furnish a robust demonstration of system suitability.

Originally, calibrator tablets were adopted to detect the influence on dissolution results due to improper alignment of the instrument, vibration in the instrument, failures in the drive chains and belts, and deaeration.[46] Thus, the calibrator tablet became an important check on operating procedures, especially in terms of consistency between laboratories on an international basis. For Apparatuses I and II, disintegrating a prednisone reference standard tablet is used to establish system suitability. USP provides a calibrator tablet of prednisone to be used to demonstrate that the bath configuration and operator practices meet specified standards. The acceptance criteria are established for each lot of calibrator tablets through collaborative studies with various laboratories around the world. Each laboratory measures the release under a defined protocol, and the results are compiled to provide the average ranges to be used as the acceptance criteria. However, the prednisone calibrator tablets have some stability issues, and the acceptance criteria for a given lot may be revised during their specified shelf life. One survey revealed that while calibrators were routinely being used to meet regulatory expectations, there were concerns that the results did not necessarily ensure accurate performance of the apparatus. Variability of the calibrator lots and the numbers of required calibrator tests were the primary concerns. Physical and mechanical performance checks have been suggested to be more important.[47]

The FDA has issued a guidance for industry, "The Use of Mechanical Calibration of Dissolution Apparatus 1 and 2—Current Good Manufacturing Practice (CGMP)," which recommends that a properly executed rigorous mechanical calibration will satisfy CGMP requirements for dissolution apparatus calibration in lieu of performance verification with specified tablets.[48] Using either Apparatus I or II, results are sensitive to the configuration of the various components such as the depth between the bottom of basket/paddle to the bottom of the vessel, centering and verticality of the shaft, leveling of the bath, and imperfections in the vessels. Bath standardization is typically monitored on a semiannual basis or after major repairs, equipment or location changes.

The ASTM, a voluntary consensus standard-setting organization, created a standard for mechanical calibration of basket and paddle dissolution apparatus building on these recommendations.[49] ASTM Standard E2503-07 provides guidance for basket and paddle dissolution apparatus setup and calibration to ensure reproducibility of results without specifying how to perform dissolution testing. This standard takes a more detailed approach to instrument setup than is currently outlined in the harmonized pharmacopeia chapters by providing quantitative criteria for shaft wobble. Shaft and vessel verticality are new parameters not currently addressed in the pharmacopoeia chapters. The ASTM standard also has tighter criteria for shaft/vessel centering (vessel offset), rotational speed, and basket wobble than that given in the Pharmacopoeia Discussion Group (PDG) harmonized dissolution testing. This change in calibration methodology, from the reliance on less stringent methodology and use of tablets to rigorous mechanical calibration, may reduce the bias and variation in measurement systems.

The Ph.Eur. recommends mechanical calibration for instrument qualification and suggests that the performance of the dissolution test instrument may be monitored by testing of an appropriate reference product.[50] The JP states that the fundamental system suitability of the dissolution apparatus must include conformance stated in its chapter of <Dissolution Test>, but specific requirements on performance verification of the apparatus are not given.[51] Hence, the International Conference on Harmonization (ICH) Topic Q4B on the harmonization of pharmacopoeial methods, in Annex 7, Dissolution Test General Chapter, notes that the harmonized dissolution test apparatus should be calibrated to ensure compliance with regional good manufacturing practice (GMP) requirements.[52]

The International Pharmaceutical Federation (FIP) recommends that the qualification of a dissolution test instrument should be performed following the calibration requirements as indicated in the FDA guidance. If additional system performance information is desired, PVT using a USP reference standard tablet or an established in-house reference product can be conducted.

13.3.4 Automation

Dissolution testing has become more valuable in development and regulation of pharmaceutical products. Obtaining the required data for the formulation screening, as well as quality control of postmarket drug products can lead to many dissolution tests in different media over long periods of time. Data in the literature showed that in the past 7 years, the total number of dissolution tests in laboratories more than doubled, from just over 2500 to more than 5800 dissolution tests.[53] A single formulation scientist can prepare enough formulations to keep two analysts busy for weeks testing the dissolution profiles. Even on short runs, the analyst must conduct many routine procedure, such as preparing dissolution media, filling the vessels, measuring the temperature of the media, dropping the testing samples, pulling aliquots of sample solution for test at specified times, and, at the end of the test, discarding the used media and cleaning the vessels for the next run. An efficient analyst in the dissolution laboratory can perform a maximum of two to four manual dissolution tests per day, depending on the run time and availability of dissolution apparatus and analytical instrument.

Automation in the pharmaceutical development laboratory offers great potential for increased efficiency and reduced laboratory errors.[54] The fully automated system consists of a dissolution tester, automated dissolution components for vessel cleaning, medium filling, tablet drops, a robotic system capable of attaching and removing USP Apparatus I baskets, a media switching valve (capable of delivering four types of media to the dissolution bath), an auto-sampler, a sample collection device, and an online UV/vis or high-performance liquid chromatography, all managed by advanced software with video recording. Fully automated dissolution tests can free the dissolution scientist from about 90% of the protocol steps in the laboratory, and has showed many other advantages. First, the automated cleaning of the bath vessels can eliminate costly analyst time and vessel breakage (safety) and also provide a more consistent cleaning. Second, except for the initial run setup, the overall automated testing can be performed without an analyst being present. Also, automated systems have capabilities to drop the tablets or capsules, record the media temperature prior to and after the test, degas and dispense the media, sample the vessels at preprogrammed times, and filter the sample solutions. The automated system reduces analyst-to-analyst variability, as well as the issues that can arise from different sampling positions, times, and tablet drops. The automated dissolution system dispenses the medium gravimetrically, which introduces less error and is done consistently from vessel to vessel and run to run. This is also important when replacing media that has been removed from sampling.

Manual dissolution bath still has an advantage for labs where the analyst does not routinely run the dissolution test or where the test methodology varies. Changes and modifications for each individual test will be easy to adapt manually.

Although fully automatic dissolution systems show advantages in terms of analyst time and reduced variability, there are several disadvantages that may need attention. The first of these is that the fully automated system is complex, with many capabilities. The problem is that the more features you have, the more errors that can be generated. The second disadvantage is the very steep learning curve for the software, which is further exacerbated if the analyst does not use the software on a regular basis. The third may be improper line cleaning. The time to troubleshoot broken seals or pumps that lost their prime and bring them back to operational status from dried-out lines, could take days. A further disadvantage is that the fully automated systems are not standardized with regard to automation when compared to different manufacturers. Operational steps that are not part of the compendial procedure should be validated. The method development and validation could be complicated and time consuming when carryover studies, line flush studies, vessel wash studies, filter studies, and sample evaporation studies all may be necessary to ensure the integrity and equivalency of the fully automatic system.

The USP proposed revision of <l092> "The Dissolution Procedure: Development and Validation" includes new sections on automation and the acceptance criteria and their interpretation.[55] This general chapter provides a comprehensive approach covering items to consider for developing and validating dissolution procedures and the accompanying analytical procedures. It addresses the use of automation throughout the test and provides guidance and criteria for validation.

13.3.5 Noncompendial dissolution methods

According to USP Chapter <1092>, a noncompendial apparatus may have some utility with proper justification, qualification, and documentation of superiority over the standard equipment. For example, a small-volume apparatus with miniature paddles and baskets may be considered for low-dosage strength products. A rotating bottle or dialysis tubes may have utility for microspheres and implants. Peak vessels and modified flow-through cells may be used for special dosage forms such as powders and stents. The

noncompendial dissolution methods have been summarized in several early publications.[56,57] Currently, conventional compendial dissolution testing remains a cornerstone of the drug development process and quality control testing. However, more studies have been done with respect to in vitro technology developments that provide the potential for formulation and analytical scientists to exceed the capabilities of the conventional dissolution test to achieve a more biorelevant testing regime.

Many current noncompendial dissolution methods have been developed based on multicompartmental models. Some of these methods not only mimic the hydrodynamics and composition of media, but also incorporate mechanical processing, digestion of real food, and gastric emptying. Examples of these systems are the TNO TIM−1 system[58] and the dynamic gastric model (DGM).[59] The TNO intestinal model (TIM) (TNO Pharma, Zeist, the Netherlands) is a computer-controlled model that simulates in vivo fluids at more realistic volumes, with more realistic fluid dynamics relative to the human stomach and small intestine. This also incorporates enzymatic activity, bile salts, and pancreatic juices.[60] DGM (Plant Bioscience Ltd, Institute of Food Research, Norwich, United Kingdom) is based on current biochemical and mechanical knowledge from the in situ stomach, including motility and shear.[61] DGM incorporates inhomogeneous mixing behavior with more realistic emptying into a model duodenum.[62] However, because of the complexity of these systems, some limitations such as result reproducibility and repeatability, cost of each test, ease of use of the method and the need of high throughput may prevent the wide use of these systems.

Some less complex systems have also been developed to simulate the physiological factors such as pH, fluid composition, and hydrodynamics to provide scientists a way of assessing product performance under more physiologically relevant conditions. These systems are typically configured to allow the transport of drug contents from a gastric compartment to a second intestinal compartment. One of the systems is the combined dissolution-absorption (permeation) model. Other studies developed a continuous dissolution-permeation system with the use of a two-compartment model consisting of two chambers (dissolution and receiver compartment) separated by a Caco-2 monolayer.[63−65] The similar design based on modification of USP apparatuses using artificial membranes instead of Caco-2 monolayers has also been reported.[66] Drug dissolution and permeation could be simultaneously determined with this model. Another example is the FloVitro dissolution system provided by the Dow Chemical company.[67] This system comprises gastric and small intestinal chambers with an additional third compartment connected to the intestinal chamber, which is intended to function as an absorptive compartment. The FloVitro approach, which integrates a tunable transfer system with deconvoluted in vivo dissolution data inputs, has the potential to combine a dynamic biorelevant dissolution test with biomodeling.

Other than multicompartmental models, dissolution apparatus with simulated GI physical forces was developed.[68,69] The novel stress test apparatus is capable of mimicking the pattern of irregular movement inside the GI tract and the physical forces applied on dosage forms during gastric emptying and intestinal transit. The system appears to have the utility for assessing the mechanical robustness of MR oral dosage formulations and their ability to withstand gastric forces and passage through the pylorus or ileocecal junction.

For these noncompendial dissolution methods, some in vitro parameters (media and hydrodynamics, physical forces, etc.) have been tested to simulate the in vivo conditions in hopes of finding successful predictions of the in vivo performance and IVIVC for oral formulations. However, despite decades of research into the area, there is not enough knowledge about these techniques, and there are very limited data available to validate the widespread use of these techniques at present. More studies are required to search for an accurate in vitro surrogate of the GI tract.

13.4 REGULATORY CONSIDERATIONS

13.4.1 The role of dissolution in product quality control

Over the past four decades, considerable scientific attention has been given to understand the mechanisms of drug dissolution and the factors (eg, formulation, manufacturing process, and physiologic factors) affecting drug dissolution, and to establish standardized methodologies for dissolution testing. The current scientific understanding and knowledge regarding drug dissolution has been predominantly based on studies of oral drug delivery systems. The FDA has published guidance documents covering the relevant topics, such as drug dissolution specifications from solid oral dosage forms and establishment of IVIVC,[2,7] demonstration of drug dissolution/release similarity when formulation and manufacturing changes are made,[5,6] and the waiver of in vivo BE studies.[4] Furthermore, the ICH Q6A guideline on establishing specifications has been developed.[70]

Because the release characteristics may have significant influence on drug BA, dissolution is an essential step in delivering drug molecules to their sites of action. Therefore, drug dissolution is a critical quality

characteristic that needs to be controlled throughout the life cycle of a product. When the hydrodynamics and media are appropriately chosen to have sufficient ruggedness and reproducibility without being over discriminating, dissolution testing can be used to ensure the batch-to-batch consistency and to detect any significant changes that could affect the in vivo performance of the product.

If IVIVC can be established and validated, then the in vitro dissolution can be a surrogate for the in vivo performance, which may be controlled in a range so that all lots with dissolution within the range are bioequivalent. Less optimal but still desirable, lots exhibiting dissolution profiles at the upper and lower dissolution limits should be bioequivalent to the clinical/BA lots or to an appropriate reference standard. In this way, the product quality is controlled with a concept of ensuring the consistency of the in vivo behavior of the product by comparing product dissolution to ensure that the product manufactured (each batch and every batch) is bioequivalent to the reference product, which in turn assures the efficacy and safety of the product. This concept leads to the basic abilities of dissolution testing: the ability to discriminate the products with quality differences, which reflect the differences of their in vivo performance and the ability to reject the batch that is not bioequivalent to the reference batch. In this chapter, the considerations for enhancing these abilities are discussed in the following sections.

Due to the in vivo implication of dissolution testing, there are significant opportunities to build upon the current regulatory decision system and to further improve its effectiveness and efficiency. These opportunities are further enhanced by the possibilities in regulatory decisions to utilize pharmaceutical development information (eg, ICH Q8[71]) and the availability of new technologies for more effective control of formulation and manufacturing variables that can affect drug dissolution when combined with a comprehensive system approach for regulatory quality assessment. The development of regulatory decision criteria based on scientific principles for quality assurance and control of drug dissolution or release rate is an important step for 21st-century pharmaceutical quality control.

Another important role of dissolution testing is in the evaluation of waiver of in vivo BA/BE studies (biowaivers), such as for preapproval and postapproval changes.[5,6,72] Provisions of biowaivers under certain conditions are provided in the Code of Federal Regulations. The regulatory basis for granting a waiver of the requirement for the submission of in vivo BA/BE data is derived from 21CFR320, where it states that either an in vitro test that has been correlated with and is predictive of human in vivo BA [320.24(b)(1)(ii)] or a currently available in vitro test that ensures adequate human in vivo BA [320.24(b)(5)] is acceptable for the evaluation of BA/BE.[73] Biowaivers based on comparative in vitro testing (eg, dissolution for solid dosage form) for different strengths of IR products can be granted under 21CFR320.22(d)(2). Recommendations on dissolution-based biowaivers for product strength equivalence for MR products is provided in the draft FDA guidance for BA and BE studies submitted in NDAs or investigational new drugs.[74] The topic of biowaivers is discussed further later in this chapter.

In the absence of IVIVC, dissolution tests can be used for product quality control to ensure that the dissolution results of the routine manufacturing batches are similar to that of the pivotal batches or BE batches. Dissolution data can be sensitive to the physical properties of ingredients, product compositions, and the manufacturing process for products; hence, dissolution tests may be able to detect batches that show significant differences in dissolution data that could lead to a significant difference in the in vivo performance of the products. Notably, the criterion of similar dissolution for quality control is limited to the same product created by the same manufacturing process. Application of the criteria to multisource products is difficult and not necessary except for some rapid-release IR products of BCS Class I and III compounds. With proper justification, different dissolution methods and specifications are acceptable for multisource products, especially for ER products. However, once the method and specifications are established for a product based on the batches that have demonstrated acceptable BA or BE, future batches of the product must meet the same specifications for dissolution.

13.4.2 Dissolution method development: regulatory considerations

Guidelines for development and validation of dissolution methods can be found in the FDA guidance and the current USP Chapter <1092>, and so they will not be discussed in detail in this chapter. To support and justify the selection of the proposed dissolution method, a report for method development should be included in the regulatory submission. The development reports generally include the information that follows in this section.

Solubility data for the drug substance over the physiologic pH range are important for dissolution method development because these data provide the basis for selection of an appropriate medium and determine whether a surfactant is needed.

A detailed description of the dissolution method being proposed for the evaluation of the product should be provided. The developmental parameters (ie, selection of the equipment/apparatus, in vitro

dissolution/release media, agitation/rotation speed, media pH, assay, sink conditions, use of sinker, and enzyme, if applicable) used to select the proposed dissolution method as the optimal test for the product are very helpful for evaluation of the dissolution method. If a surfactant is used, data should be included to support the selection of the type and amount/concentration of the surfactant. The conditions used for each test should be clearly specified. The dissolution profile should be complete and cover at least 85% of drug release of the label amount or whenever a plateau (ie, no increase over three consecutive time-points) is reached. A systematic approach for selection of the dissolution conditions is to use design of experiments (DOEs), which considers various factors simultaneously, not only for the main effects of the individual factors, but also for the interactions among the factors. For example, in Fig.13.5, the results of dissolution at 15 minutes of a model product are shown for an example DOE by varying media (pH 1, 4.5, 6.8, and water), surfactant [cetyl triethylammonium bromide (CTAB), lauryl dimethylamine-oxide (LDAO), and benzalkonium chloride (BenzChlor)], concentration of surfactant (0.5% and 1%) and paddle rotation speed (50 and 75 rpm). Here, because the complete dissolution was not obtained, the media at pH 4.5 or 6.8 and water were not appropriate. Based on the physicochemical properties of the compound, through data analysis and further experiment, pH 1 media at paddle rotation speed of 50 rpm with surfactant CTAB at concentrations of 0.5% were selected.

In addition to the factors listed here, ionic strength in the medium may play a role (Fig. 13.6, where phosphate buffer is used). Although the pHs for all three profiles generated are at 7.5, the ionic strengths of potassium phosphate are different (0.05, 0.1, and 0.2 M, respectively), which make differences in the dissolution of the model product.

For comparison, at least twelve samples per test are recommended. In general, the following sampling time points are selected: 10, 15, 20, 30, 45, 60, 90, and 120 minutes for IR formulations, and 1, 2, 4, 8, 12, and 24 hours for ER formulations.

Complete dissolution profile data should be provided, including individual, mean, and SD (or %CV) at each time point, as well as profiles for the product. The dissolution data at each time point are usually reported as the cumulative percentage of drug dissolved (the percentage is based on the product's label claim).

Data to support the discriminating ability of the selected dissolution method are critical for evaluation. In general, the testing conducted to demonstrate the discriminating ability of the selected dissolution method should compare the dissolution profiles of the reference (target) product and the test products that are intentionally manufactured with meaningful variations for the most relevant critical manufacturing variables (ie, ±10−20% change to the specification-ranges of these variables) including critical material attributes (CMAs) and critical process parameters (CPPs). Fig. 13.7 shows the dissolution profiles for two batches

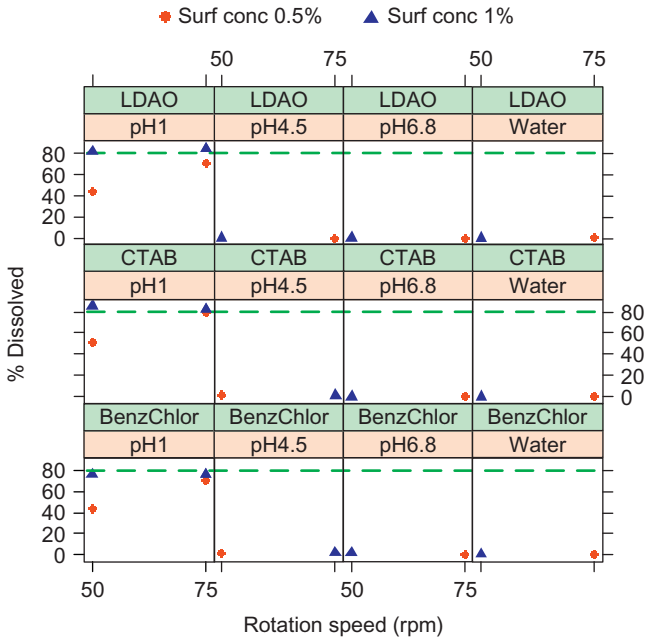


FIGURE 13.5 Example DOE for effects of media pH, surfactant type and concentration, and rotation speed upon dissolution.

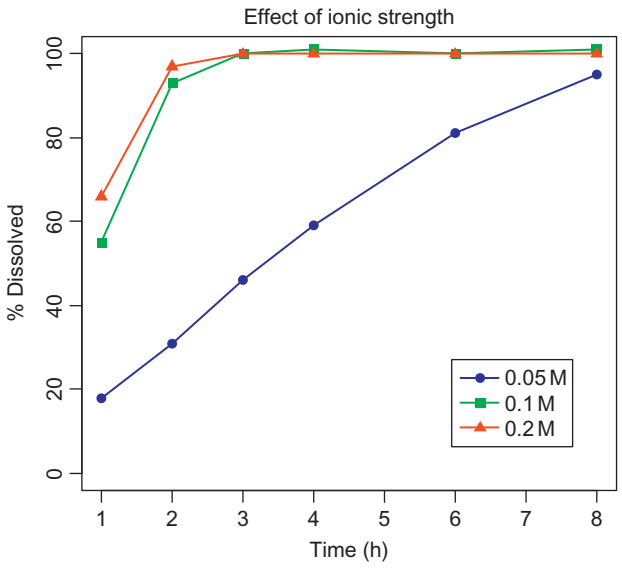


FIGURE 13.6 Effects of ionic strength on the dissolution of a model product.

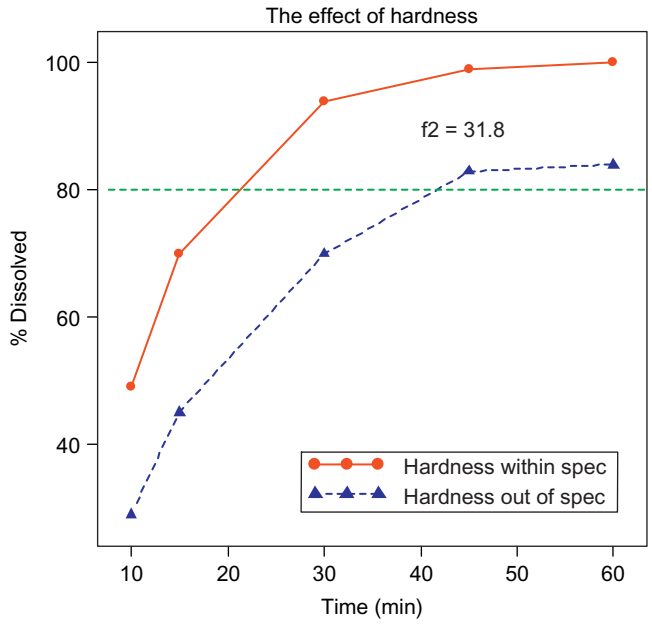


FIGURE 13.7 Effect of tablet hardness on dissolution.

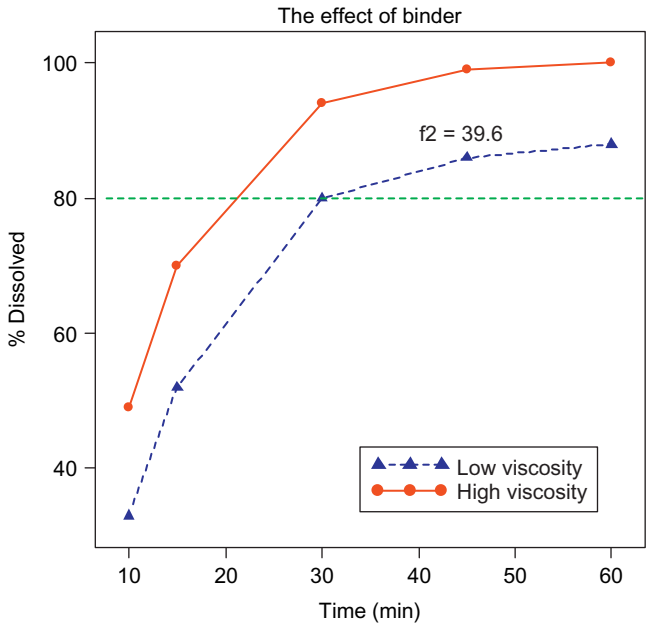


FIGURE 13.8 Effect of binder viscosity on the product dissolution.

of tablets with different hardness (by varying compression force, a CPP). As shown, the release of the batch represented by the line with up triangles is much slower, compared to the batch represented by the line with solid circles, and the former represents an incomplete dissolution although the assay results show the comparable potency between the two batches. This demonstrates that the dissolution method is able to distinguish these two batches with an f2 (similarity factor used to compare two dissolution profiles) value of 31.8, which is much less than 50, indicating that the significant difference between the two batches is detected by the dissolution method. Fig. 13.8 shows the dissolution profile comparison between a batch with the viscosity of the binder, a CMA, within the specification and a batch with the viscosity out of specifications. The dissolution method is able to distinguish these two batches with an f2 value of 39.6 indicating significant difference between the two batches. These two sets of data support the discriminating ability of the dissolution methodology selected.

Further, providing data that show the selected dissolution method is able to reject batches that are not bioequivalent to the reference batches is helpful. Fig. 13.9 shows the dissolution profiles for two batches A and B. A is the batch used in the pivotal clinical studies, which has demonstrated efficacy and safety. B is a batch with manufacturing changes and has been shown not bioequivalent to A. As seen, the dissolution is able to pick up the differences in this case. When setting appropriate acceptance criterion, these two

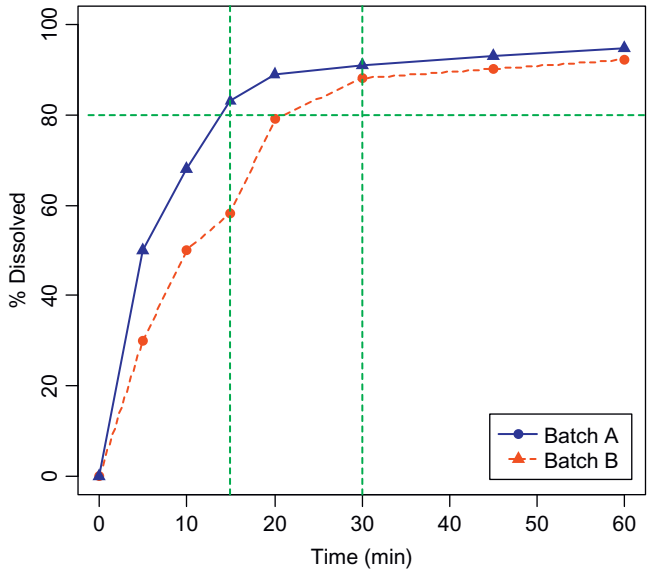


FIGURE 13.9 Dissolution profiles of two batches (A and B), which are not bioequivalent.

products can be distinguished and the product with poor quality (not bioequivalent to the clinical batch) can be rejected. Another example is shown in Fig. 13.10, where the dissolution shows distinctions of profile shapes, indicating the differences in clinical performance. The line with red empty circle has an initial burst, which is higher than the blue line with solid circle, followed by a slowdown and an increase at a much later time compared to the profile with solid

circle. The slowdown of the release may be responsible for the ineffectiveness of the product.

On the other hand, dissolution methods may sometimes overdiscriminate the formulations as indicated in Fig. 13.11, which shows the in vitro and in vivo profiles for the target formulation and three variants. Although the in vitro dissolution profiles show significant differences among target formulation and variants, all the in vivo profiles are bioequivalent. To make an appropriate judgment using all available information, knowledge management plays a critical role.

In addition, the validation data for the dissolution method (eg, system suitability test, method robustness, deaeration if necessary, and validation between manual and automated procedures) and analytical method (ie, specificity, precision, accuracy, linearity, stability, etc.) are needed to support the proposed method.

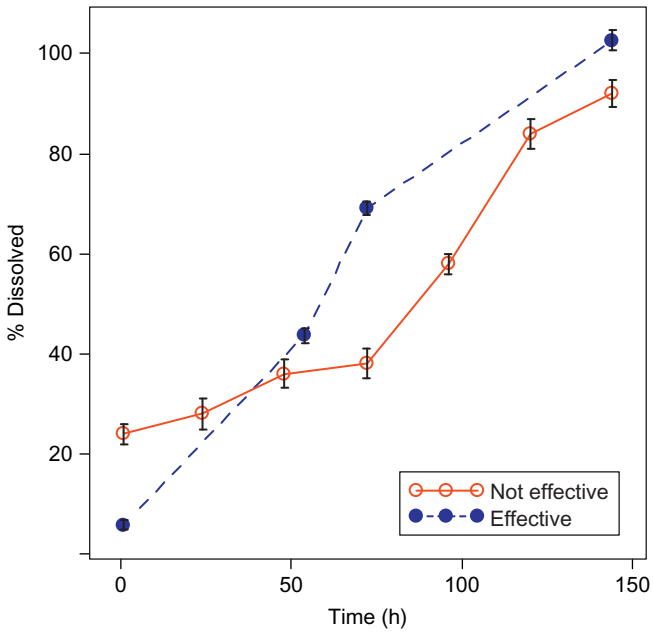


FIGURE 13.10 Dissolution profiles with different shapes have clinical efficacy differences.

13.4.3 Setting regulatory acceptance criteria for dissolution testing

When setting the dissolution acceptance criteria, the clinical performance of the product is the first and most important consideration. Thus, the data of the dissolution profile (eg, 10, 15, 20, 30, 45, 60, 90, and 120 minutes for IR products; 0.5, 1, 2, 4, 8, 12, and 24 hours for ER products; $n = 12$) from the pivotal clinical batches should be used for setting the dissolution acceptance criteria of the product (ie, specification-sampling time point and specification value). The manufacturing information of those batches (number, date, site, and size of the batch made) should also be provided. Since the pivotal clinical studies may last for a period of time due to patient enrollment and other factors, clinical supplies may not always be the freshest batches. In this case, the stability of the batches may be considered when setting the acceptance criteria.

Because the batches participating in pivotal clinical trials are limited, the dissolution range, within which the products will be bioequivalent to the reference

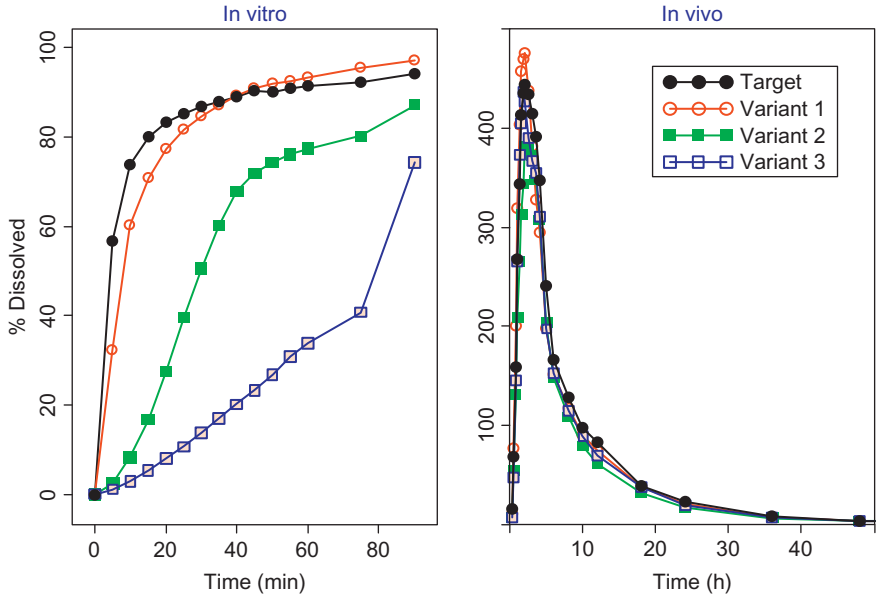


FIGURE 13.11 Overdiscriminating dissolution method.

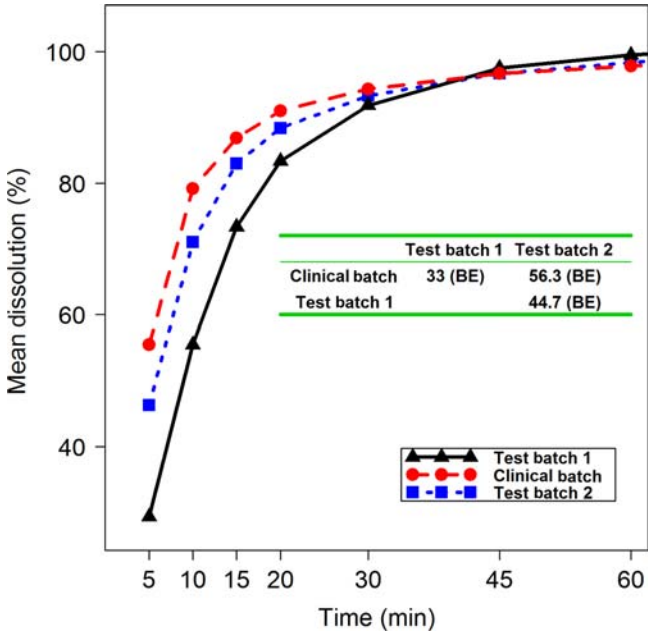


FIGURE 13.12 The dissolution profiles of three bioequivalent batches of formulations.

batch, can be difficult to identify. Further, the limited batches may not represent the real variability of the product caused by many factors, including the materials and the processes. In these ways, taking the batches involved in all the relevant clinical studies into consideration would be very helpful. The batches, which are bioequivalent to the reference batch, could be used for defining the range of the acceptance criteria. In the example shown in Fig. 13.12, the dissolution of the clinical batch (the dashed line with solid circles) is the highest compared to the other two batches: Test Batch 1 (solid line with up triangles) and Test Batch 2 (dotted line with solid squares). The dissolution profiles are similar between the clinical batch and Test 2 (f2: 56.3), but comparisons between the clinical batch and Test 1 (f2: 33) and between Test 1 and Test 2 (f2: 44.7) do not show similarity. On the other hand, however, a clinical study showed BE among the three formulations, with Test Batch 1 barely passing the BE criteria compared with the clinical batch. Therefore, the dissolution of Test Batch 1 could serve as the lower limit when setting the dissolution criterion. Another example is shown in Fig. 13.9 in the last section, when the acceptance criterion is set as $Q = 80\%$ at 30 minutes, both batches A and B would pass although they are not bioequivalent. Therefore, the criterion should be set as $Q = 80\%$ at 15 minutes, which would be able to differentiate these two batches and reject the one that is not bioequivalent to the clinical batch.

The in vitro dissolution profile should encompass the timeframe over which at least 85% of the drug is dissolved. If incomplete dissolution is occurring, the profile should reach the plateau (ie, no significant change in three consecutive time points).

For IR products, the selection of the specification time point should be where $Q = 80\%$ dissolution occurs. However, for a slowly dissolving product or a poorly water-soluble drug, specifications at two time points may be adequate. The first time point should include a dissolution range (eg, 40–60% dissolution), and the second time point should be where $Q = 80\%$ dissolution occurs.

For ER products, the establishment of at least three specification time-points covering the initial, middle, and terminal phases of the complete dissolution profile should be set. The specification ranges should be based on the overall dissolution data generated at these time points. If the maximum amount released is less than 80%, the last time point should be the time when the plateau of the release profile has been reached. In general, the selection of the dissolution specification ranges is based on mean target value $\pm 10\%$ and not less than (NLT) 80% for the last specification time-point. Wider specification ranges may be acceptable if they are supported by a validated IVIVC model.

Delayed release (enteric coated) products should have the following specifications established in both the acid stage and buffer stage per USP:

- Acid stage: No individual tablet exceeds 10% dissolved at 2 hours.
- Buffer stage: The selection of the specification time point should be where $Q = 80\%$ dissolution/release occurs. If it is a delayed release (DR)/ER product, at least three specification time-points covering the initial, middle, and terminal phases of the complete dissolution profile should be set at the buffer stage based on the ER product listed.

The acceptance criteria is recommended to be set based on the average in vitro dissolution data equivalent to USP Stage 2 testing ($n = 12$) for each lot under study for a tighter control. To illustrate this and provide the quantitative basis to set the dissolution acceptance criterion for an IR solid oral drug product, a simulation study of the dissolution profiles was conducted on a batch of a solid oral dosage form with 1 million units (tablets). The desired acceptance criterion was $Q = 80\%$ of label claim at 30 minutes. A total of 110 scenarios were examined, which include various combinations of two determinants: batch mean and SD values of dissolution at 30 minutes. The mean of the dissolution of the 1 million tablets varied from 79% to 89% of the label claim with 1% increment (11 levels) and the SD varied from 1% to 10% of the label claim with 1% increment (10 levels). For each of the 110 scenarios, 10,000 replicates were simulated. Based on USP

Chapter <711>, three stages of dissolution testing were conducted. The failure index at each stage for each scenario was calculated as the percentage of failed replicates in the 10,000 trials. The simulation was implemented by a function written in R software (version 2.15.3, R foundation for statistical computing). A contour plot was generated for each stage to demonstrate the relationship between the failure index and the two determinants (dissolution mean and SD).

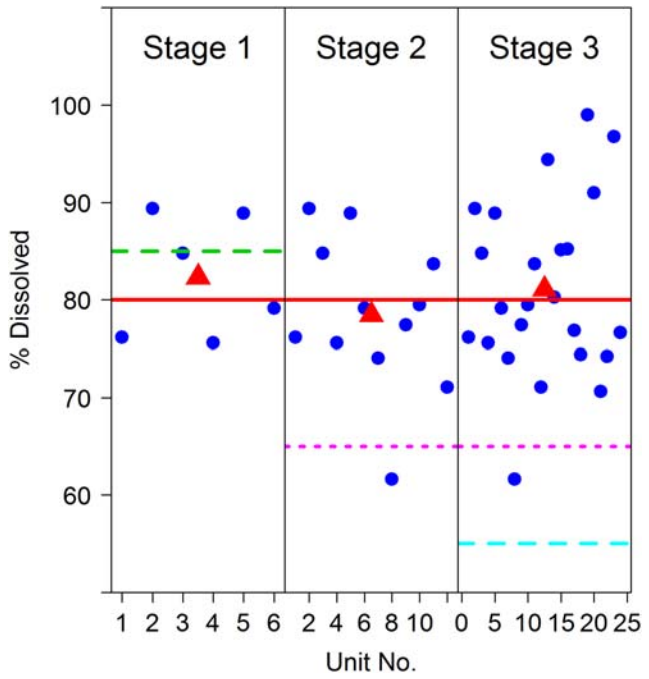


FIGURE 13.13 A typical case of three-stage dissolution for a time point at which the desired acceptance criterion is $Q = 80\%$ for an IR oral solid dosage form. The solid circle represents the dissolution value for individual unit, and the solid triangle is the sample mean. This batch does not pass Stage 1, which requires every unit ≥85% (the horizontal dashed line shown in Stage 1). The batch does not pass Stage 2 because the sample mean <80% (the solid horizontal line), and there is a unit less than 65% (the dotted horizontal line). It passes Stage 3 since the mean is above 80%, not more than two units less than 65% (only one less than 65%) and no one unit is less than 55% (the dashed horizontal line shown in Stage 3).

A typical case of three stages of dissolution testing at a given combination of mean and SD is shown in Fig. 13.13. The solid circles represent the individual dissolution values and the solid triangles are the mean values.

The failure index for each stage is dependent on the mean and SD values, as shown in Fig. 13.14. The contour lines represent the failure index at a given combination of the mean and SD values. As shown, in order for a batch to have 80% chance to pass Stage 1, the mean dissolution value of this batch must be at least 87% with very low SD of less than 1%. If the SD increases to 2%, the mean value must be at least 89%.

On the other hand, simulations indicate that if the mean dissolution value is 81%, the drug product's batch would have 80% chance to pass Stage 2 when the SD is less than 2.5%. Even when the SD increases to 8% (10% CV), the batch would have 80% chance to pass Stage 2 if the dissolution mean value is >84.5%. These observations by simulation indicate significant differences of passing the specification between Stage 1 and Stage 2 (or 3). Due to this difference, the acceptance criterion would be very permissive for Stage 2 and 3, if the setting of the criterion is based on Stage 1 dissolution data. Therefore, this suggests setting of the dissolution acceptance criterion be based on Stage 2 dissolution data is advisable. The 20% failure index contour line for Stage 2 (middle panel in Fig. 13.14) provides a quantitative estimate based on the observed mean and SD data and could serve as the guideline for the setting of the acceptance criterion for the dissolution test of IR solid oral dosage forms.

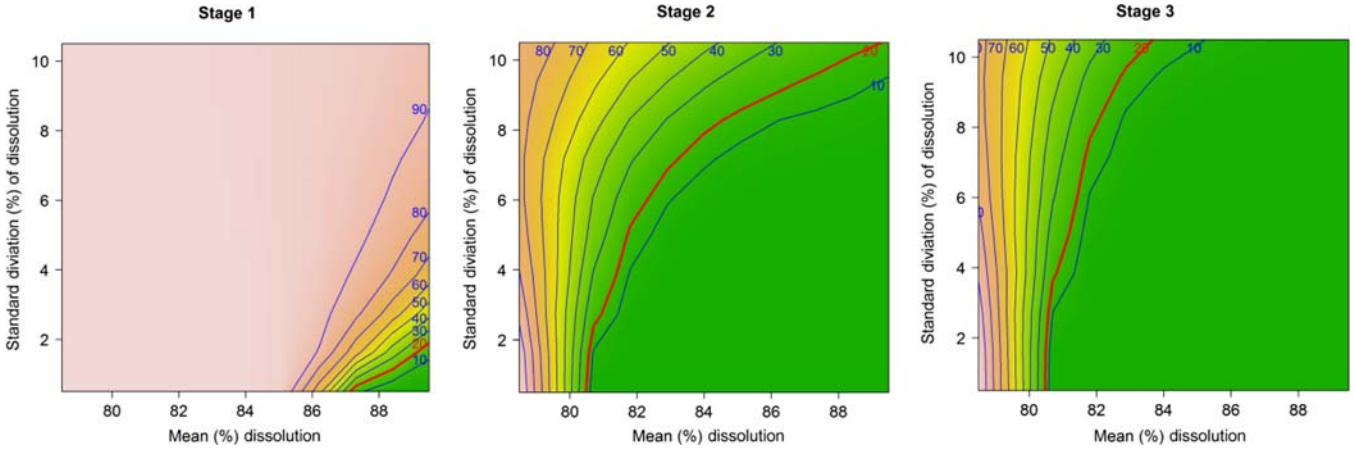


FIGURE 13.14 The influences of mean and variability of dissolution data on the failure rates for three stages of dissolution testing of IR formulation by simulation.

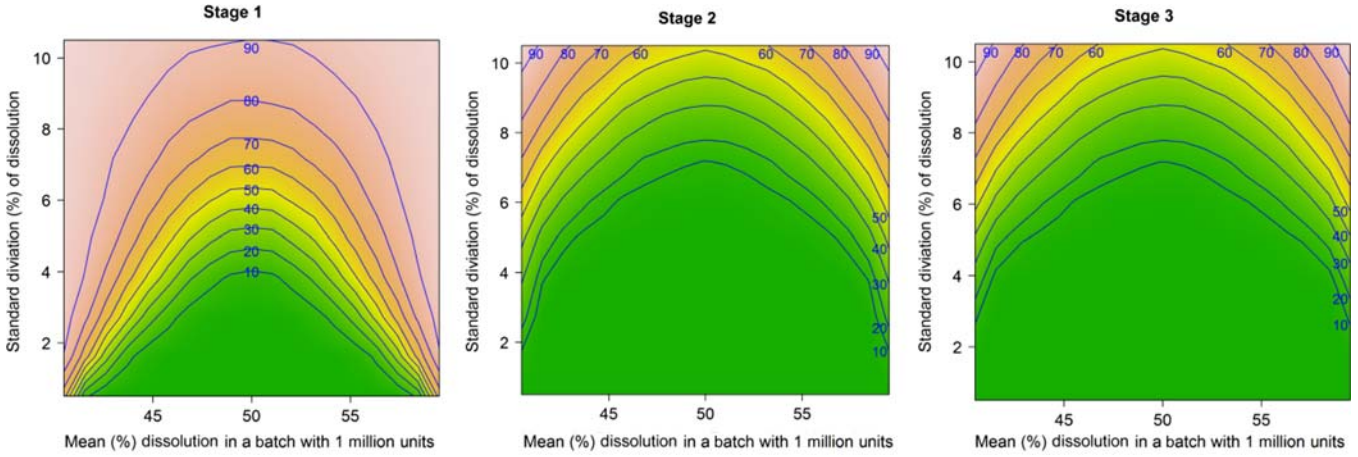


FIGURE 13.15 The influences of mean and variability of dissolution data on failure rates for three stages of dissolution testing of ER formulations by simulation.

This study provides scientific support for the current FDA's practice/recommendations on using Stage 2 data for the setting of the dissolution acceptance criterion. In addition, the study provides an example of a quantitative way to select the acceptance criterion's sampling time point based on the mean and SD of the dissolution data according to Stage 2 contour line with a 20% failure index in Fig. 13.14.

A similar simulation study was conducted to estimate the three stages of the dissolution for ER formulation, as shown in Fig. 13.15, and similar conclusions can be made, although the distinction between Stage 1 and Stage 2 (or 3) for ER formulation is not as dramatic as that for IR formulation.

Notably, the 80% success rate or the 20% failure rate discussed here for the simulation is for illustration purposes, to indicate the statistical power. One could argue that a failure rate of 20% is too high in manufacturing and should not be used as the objective to set the specifications for dissolution. However, the information from the simulation is valuable to assist setting a proper specification based on the clinical performance of the clinical batches using a quality-by-design approach to ensure the consistent performance of product by controlling the lot-to-lot variation, and to prevent the release of any lots with dissolution profiles that have been shown not bioequivalent in clinical study.

As discussed earlier, if a valid IVIVC model has been established, the model should be used to set the specifications for dissolution. Specifications should be set for the product with the fastest and slowest release rates to result in a maximal difference of 20% in the predicted C_{max} and area under the curve (AUC).

If there is no IVIVC/IVIVR established, the following steps may be considered to set up the dissolution acceptance criteria. The procedure fixes the Q value at 80% for IR products and presents a typical case for ER products using clinical batches as the bases for consideration.

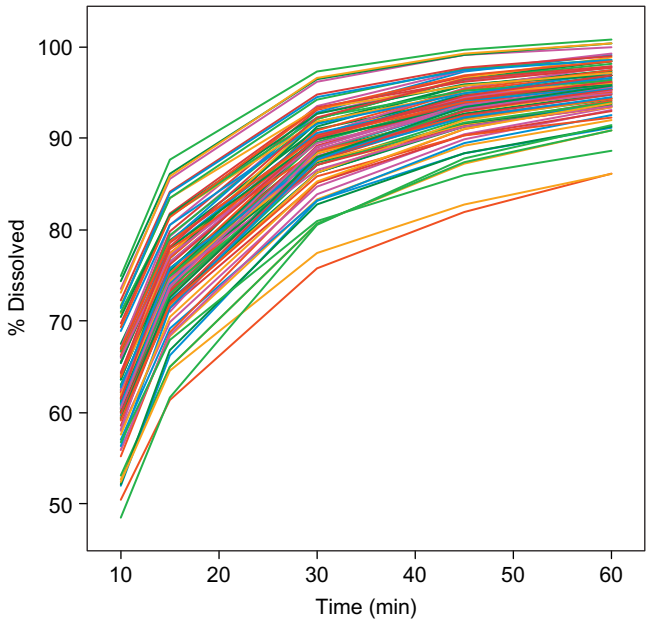


FIGURE 13.16 A plot of the dissolution profiles of all available batches for an IR product.

Step 1: Plot the available dissolution profiles to appreciate the variability and select the potential specification time points. Find the root causes if the variability is too high (>10% CV). Resolve the quality problems in manufacturing, assay, or both.

For IR formulation, the time point can be chosen where $Q = 80\%$ dissolution occurs. For ER formulation, three time points should be selected at initial, middle, and terminal phases of the complete dissolution profile. As shown in Fig. 13.16, in the current example, a

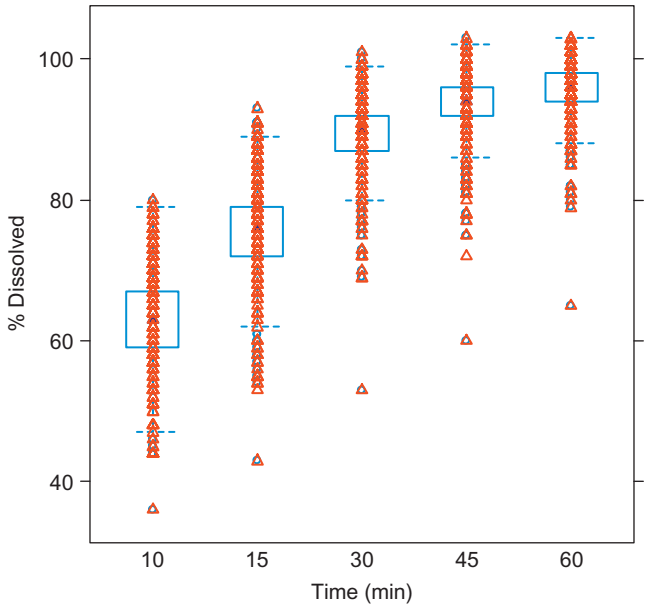


FIGURE 13.17 Box plots for the dissolution data at each time point. The mean, 25%, and 75% quartiles are identified for the distribution for an IR product.

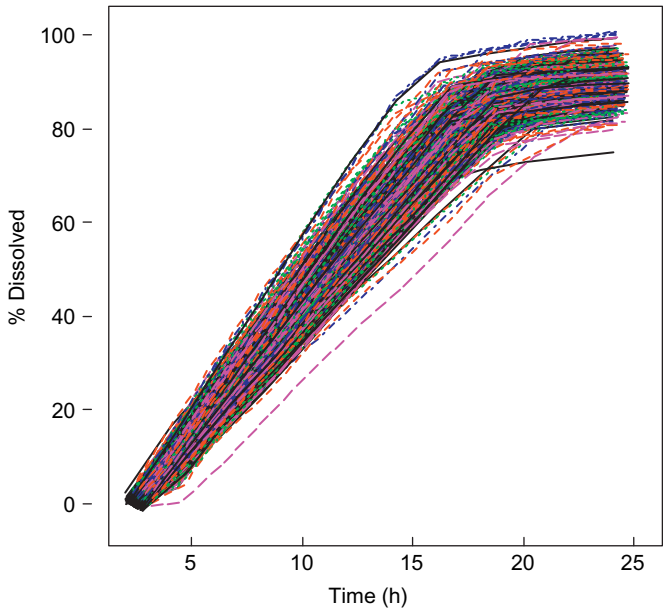


FIGURE 13.18 A plot of the dissolution profiles of all available batches for an ER product.

total of 90 profiles available for an IR product are plotted to get an insight regarding the general shape of the dissolution profile and the variability. Pay attention to the outliers and special features, such as strength-dependent behavior. To further appreciate the variability, Fig. 13.17 uses box plots for about 1000 individual units to identify the mean, the 25th, and the 75th quartiles of the distribution at each time point. From the plot, two observations can be made for this product. First, the variability is relatively large, which could raise a concern regarding the product quality, and second, the 30-minute time point could be a potential specification sampling time for this IR product. If the large variability is caused by product quality, the sponsor should make every effort to control the CMAs and CPPs. If the root cause for variability is assay, improvements of analytical procedure are needed. Before setting the appropriate dissolution specification, the variability should be reduced to a reasonable level (CV $<$ 10%).

For ER products, the same procedure can be followed. As shown in Figs. 13.18 and 13.19, the dissolution profiles and the distribution of all individual units at each time point are plotted. From these observations, potential specification time points at 6, 10, and 18 hours are selected.

Step 2: Plot the clinical batch dissolution data at the potential specification time point.

As mentioned previously, the dissolution of the clinical batches, which have shown the acceptable efficacy and safety, are the standards for setting dissolution

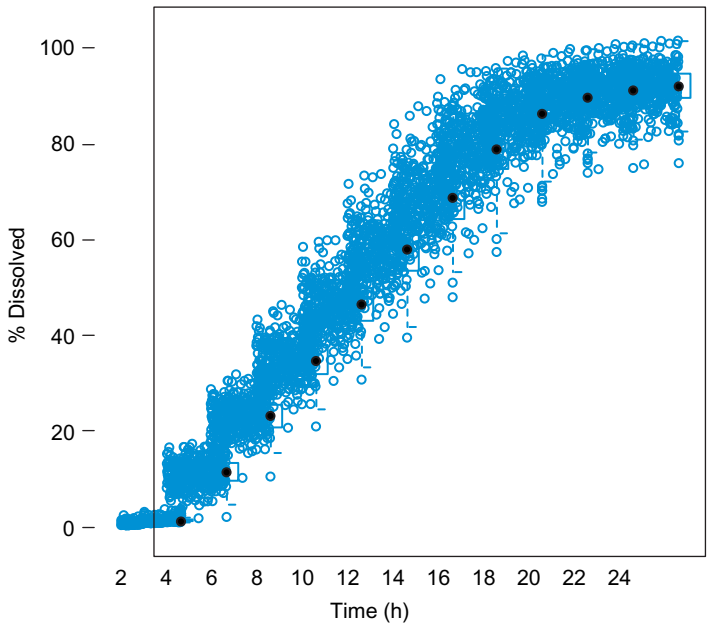


FIGURE 13.19 Box plots for the dissolution data at each time point. The mean, 25%, and 75% quartiles are identified for the distribution for an ER product.

acceptance criteria. As shown in Fig. 13.20, the dissolution of individual unit at 30 minutes for 26 batches used in pivotal clinical trials (showed acceptable efficacy and safety), are examined with an intention to set the acceptance criterion as $Q = 80\%$ at 30 minutes. As seen, although some batches could not pass Stage 1 due to individual unit values less than

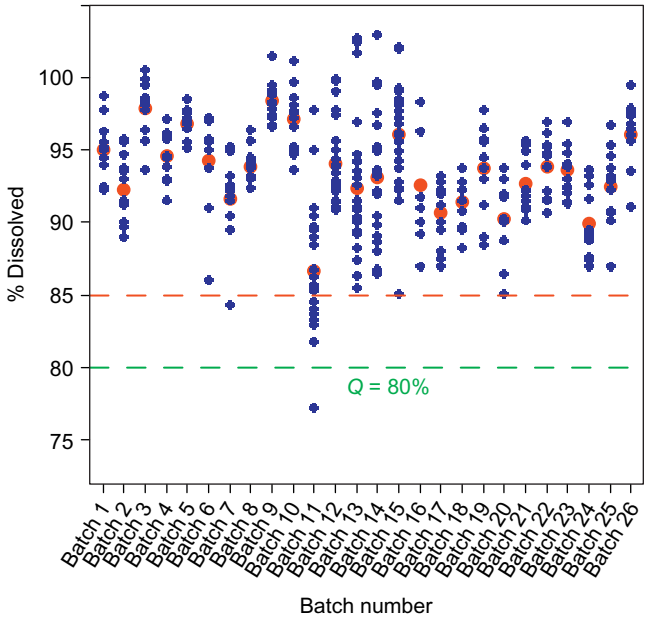


FIGURE 13.20 An example case of dissolution performance of 26 clinical batches of an IR formulation at the selected potential specification time point.

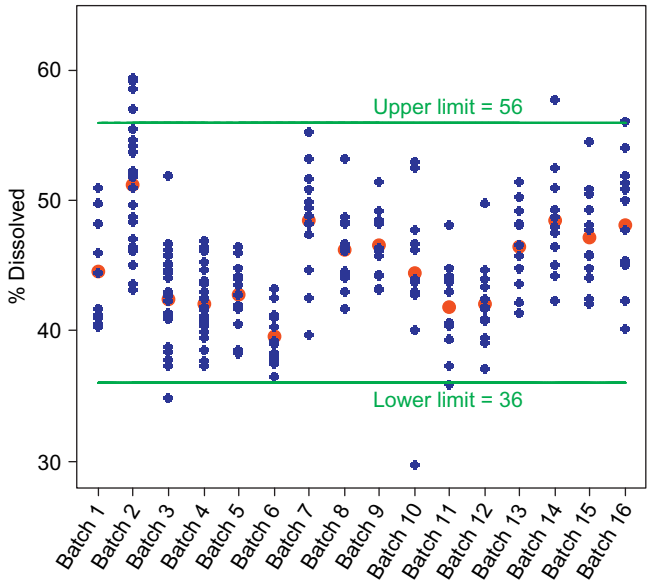


FIGURE 13.21 An example case of dissolution performance of 16 clinical batches of an ER formulation at one of the potential specification time points.

85% (Stage 1 requires all units to be $\geq Q + 5\%$), all batches will pass Stage 2 (the mean, represented by the bigger red dots (black in print versions) are above 80% and no single units below 65%). Therefore, the potential specification time point (30 minutes in this example) seems adequate.

Similarly, Fig. 13.21 shows an example of one of the specification time points (at 10 hours) for an ER product. As seen, the acceptance criteria 36–56% are based on the totality of dissolution data for 16 batches used in the pivotal clinical studies, considering both the mean (larger red dots (black in print versions)) and individual values (blue dots (gray in print versions)). Although some batches would not pass Stage 1, which requires all units to be within the specified range, all batches will pass Stage 2 since the batch means are within the specified range and no single unit out of the range by more than 10%.

Step 3: Confirm the selected specification time points by inspection of available dissolution data, including stability data.

When confirming the selection of the specification time point and the acceptance criteria, the simulation results shown in Figs. 13.14 and 13.15 should be considered to see if the specification will allow the product to pass the acceptance criteria with high power based on the middle panel in each graph. Generally, a bootstrap analysis for the clinical batches may be performed to get the mean and SD values, which can then be used to check against the contour line for the failure index.

Obviously, one should check the dissolution specification using the dissolution performance of the product employed in the relevant clinical studies, successful or not, which may provide further understanding of the product performance and its clinical relevance.

As mentioned previously, stability data on dissolution are important and should be taken into account when setting dissolution specifications. While significant changes in dissolution over the course of a stability study are not acceptable, minor changes in dissolution of an aged product may occur. In addition, the dissolution performance during stability for hard and soft gelatin capsules and gelatin-coated tablets due to crosslinking should be noted.[75]

If any issue arises during confirmation, go back to Step 2. This iteration would provide an optimal selection for dissolution specifications.

If the dissolution is not a sensitive indicator for the product quality compared to disintegration testing, it may be substituted by disintegration, as indicated by ICH Q6A.[70]

Note that the final determination on the acceptability of the dissolution acceptance criteria is made by regulatory review. The adequacy of the proposed dissolution acceptance criterion for the product is made during the NDA review based on the totality of dissolution data provided in the application.

The general approach used in setting dissolution specifications for NDAs are summarized in Fig. 13.22.

A dissolution test is commonly used as a quality control tool for ANDAs in the generic industry. For

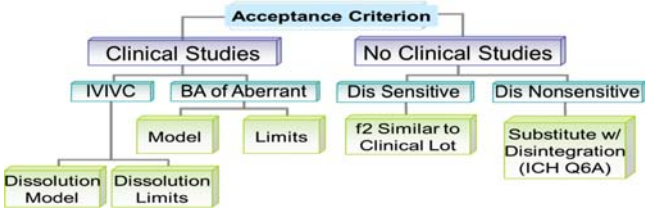


FIGURE 13.22 General scheme for determining the dissolution acceptance criteria for NDAs.

this purpose, the dissolution method and accompanying specifications are established based on the dissolution performance of the pivotal BE batch that has been demonstrated to be bioequivalent to the referenced listed drug (RLD) and the stability batches of the same formulation. Except for some rapidly dissolving BCS Class 1 or 3 drug products, for which waiver of in vivo BE studies may be granted under certain conditions (as discussed in the next section), IVIVC is very limited for ANDAs at present. In addition, the total number of batches of acceptable product available for setting dissolution specification for an ANDA is less than that of an NDA. Therefore, the applicability of the steps outlined here for NDAs may be limited for ANDAs. However, the concept is still valid for consideration in setting the specifications for ANDAs to ensure that the dissolution test has both ruggedness and repeatability for routine operation and, ideally, to have sufficient sensitivity to identify the differences that could potentially lead to a significant change in the in vivo performance of the products.

In many cases, the dissolution methods listed in the FDA's dissolution database[76] and in the USP are used for new ANDAs. Due to the lack of valid IVIVCs for multisource products and the fact that the dissolution of a given drug product can be affected by many factors, including formulation, manufacturing process, and powder characteristics such as particle size distribution and polymorphism of APIs and excipients for the product, the methods in the FDA database or USP may or may not be suitable for new ANDA products that have been shown to be bioequivalent to the RLDs. Therefore, ANDA sponsors may develop new dissolution methods/specifications if the existing methods in the FDA database or USP are not adequate for a new generic product. The FDA perspective and decision trees for selection of dissolution test and the suggestions for submission of dissolution data for ANDAs of both IR and ER products can be found in the literature.[77] Similarly, USP may include multiple dissolution tests or drug release tests in a same monograph for multisourced products that have different in vitro dissolution behavior but yet are bioequivalent to the RLD, which is common for ER products. However, it should be noted that, in some cases, only some the products meeting the different requirements stated in a same monograph are bioequivalent to the RLD. Some of the pharmacopoeial monographs may cover products with different therapeutic equivalence (TE) codes or for different uses, such as for twice-a-day or for once-a-day dosing; hence, those products are not expected to be bioequivalent. Product TE codes can be found in the FDA Orange Book at http://www.accessdata.fda.gov/scripts/cder/ob/default.cfm.

13.4.4 Biowaiver considerations and comparison of dissolution profiles

Waiver of BE studies in human (biowaivers) may be possible for (1) certain types of generic drug applications (ANDAs), (2) approval of lower strengths (or higher strengths in some cases) of products based on the successful BE of the designated strength of the product, and (3) postapproval changes to NDAs or ANDAs.

For ANDAs, biowaivers may be possible based on the BCS and the formulation and dissolution of products.[4] According to the FDA draft guidance,[5,6] biowaivers can be justified for rapidly releasing IR products containing BCS Class 1 drug substances (highly soluble and highly permeable), as well as BCS Class 3 drug substances (highly soluble and low-permeable). Rapid dissolution is defined as not less than 85% of the label amount of the drug substance dissolves within 30 minutes using USP Apparatus I at 100 rpm, or USP Apparatus II at 50 rpm (or 75 rpm with justifications), in a volume of 500 mL or less of each of the following media: (1) 0.1 N HCl or simulated gastric fluid USP without enzymes, (2) a pH 4.5 buffer, and (3) a pH 6.8 buffer or simulated intestinal fluid USP without enzymes. Dissolution profiles of the generic drug products under these test conditions should also be similar to that of the RLDs. Class 1 drug products should not contain any excipients that will affect the rate or extent of absorption of the drug. For Class 3 drug products, additional requirements are imposed for biowaivers, which includes (1) the drug product (test and reference) is very rapidly dissolving (NLT 85% dissolves within 15 minutes) in each of the recommended media and (2) the test product formulation is qualitatively the same and quantitatively very similar (eg, falls within SUPAC IR level 1 and 2 changes in composition) to the RLD.

Waivers of in vivo BE studies may be possible for the approval of lower or higher strengths of a product based on the acceptable BA/bioequivalency of the designated BE strengths of the product with the additional

requirement that (1) the formulations of the designated BE strengths of the test and the RLD products are proportional to their corresponding strengths of products and (2) dissolution testing of all strengths is acceptable (similar).[78]

Changes in SUPAC for IR and MR solid oral dosage forms in different classes, such as components, manufacturing site, scales, equipment, and process, can be classified into three levels.[5,6] Among the changes, Level 2 and Level 3 are the cases where additional dissolution test is needed to justify the changes and to ensure that the product quality and performance of the product remain similar upon the SUPAC changes, and to reduce the burden for the regulatory agency and the industry. One of the common criteria used for comparison of dissolution profiles is the use of similarity factor f2 (defined in Eq. (13.1)) for justification of the similarity of dissolution performance of the test (or postchange) and the reference (or prechange) batches. When f2>50, it can be considered that dissolution is similar for the prechange and postchange batches. Dissolution profile comparisons play an important role under these conditions. According to the SUPAC guidance,[5,6] an in vivo BE study is needed when the changes are likely to have significant impact on the product quality and efficacy. The BE study may be waived if IVIVC has been established for the product. The development of IVIVC is discussed in chapter "In Vitro/In Vivo Correlations: Fundamentals, Development Considerations, and Applications."

As discussed previously, the comparison of in vitro dissolution profiles is immensely important for justification of biowaivers and for SUPACs. The mathematical approaches for quantitative comparison of dissolution profiles are described in the literature. One of these approaches was published by Moore and Flanner.[79] They described a "similarity factor," also known as the *f2 index*, which is a logarithmic transformation of the average of the squared vertical distances between the test and reference mean dissolution values at each dissolution time point, as shown in Eq. (13.1):

$$f2 = 50 \times \mathrm{Log}\left\{\left[1+\frac{1}{n}\sum_{t=1}^{n}(R_t-T_t)^2\right]^{-0.5}\times 100\right\} \tag{13.1}$$

where Log = logarithm to base 10, n = number of sampling time points, $\sum$ = summation over all time points, and R_t and T_t = the reference and test dissolution values (mean of at least 12 dosage units) at time point t.

The value of f2 is 100 when the test and reference mean profiles are identical. The maximum distance between mean dissolution profiles at any time point cannot exceed 100%, in which case the value of f2 will be close to zero. An average difference of 10% at all measured time points results in a f2 value of about 50. Therefore, the values of f2 between 50 and 100 reflect less than 10% overall difference between the profiles suggesting similarity of the two dissolution profiles and thus, the release performance of the two products.[2,4–7]

The f2 index is used frequently by the pharmaceutical industry to compare mean dissolution profiles. Different FDA guidance documents suggest the use of the f2 index.[2,4–7] The f2 index provides an easily calculated single number to describe the comparison of dissolution profiles. However, the index has certain disadvantages.[80–83] For example, Eq. (13.1) does not take into account the variability or correlation structure or direction of the data. Also, the values of f2 are sensitive to the number of dissolution time points. Finally, the basis for the criteria to determine the similarity between dissolution profiles is somewhat ambiguous.

There are some restrictions for the f2 index, including the fact that the variability of the dissolution data should be less than 20% at early time points and less than 10% at other time points, and that only one time point with more than 85% dissolution should be used in the calculation. However, these restrictions were sometimes violated.

An important issue in the dissolution profile comparison using f2 is the practical significance of differences between mean dissolution profiles. That is, how large can the difference between mean dissolution profiles be before the differences are likely to affect in vivo performance?

Knowing how consistent f2 similarity is with the criteria for BE is important for assuring similarity in product performance. For this purpose, addressing the following two questions is necessary: (1) how likely the two products determined to be similar in vitro are not bioequivalent in vivo (false positive) and (2) how likely the two products determined to be dissimilar in vitro are bioequivalent in vivo (false negative).

A study was conducted to investigate the consistency between the in vitro dissolution profile comparisons using an f2 matrix and in vivo BE using the 80–125% criteria for an ER formulation.[84] The study utilized a simulation approach to examine several potential scenarios to get a general picture about this issue. Fig. 13.23 exemplifies a set of simulated dissolution profiles.

In the figure, the red (gray in print versions) solid lines in each panel are the dissolution profiles for the reference product, while the blue (black in print versions) lines are those for the test product. As can be seen, when the parameters of the test and reference used for simulation are the same, the profiles of the reference and the test are exactly overlapped (panel 25 with f2 = 100). However, as the parameters of the test

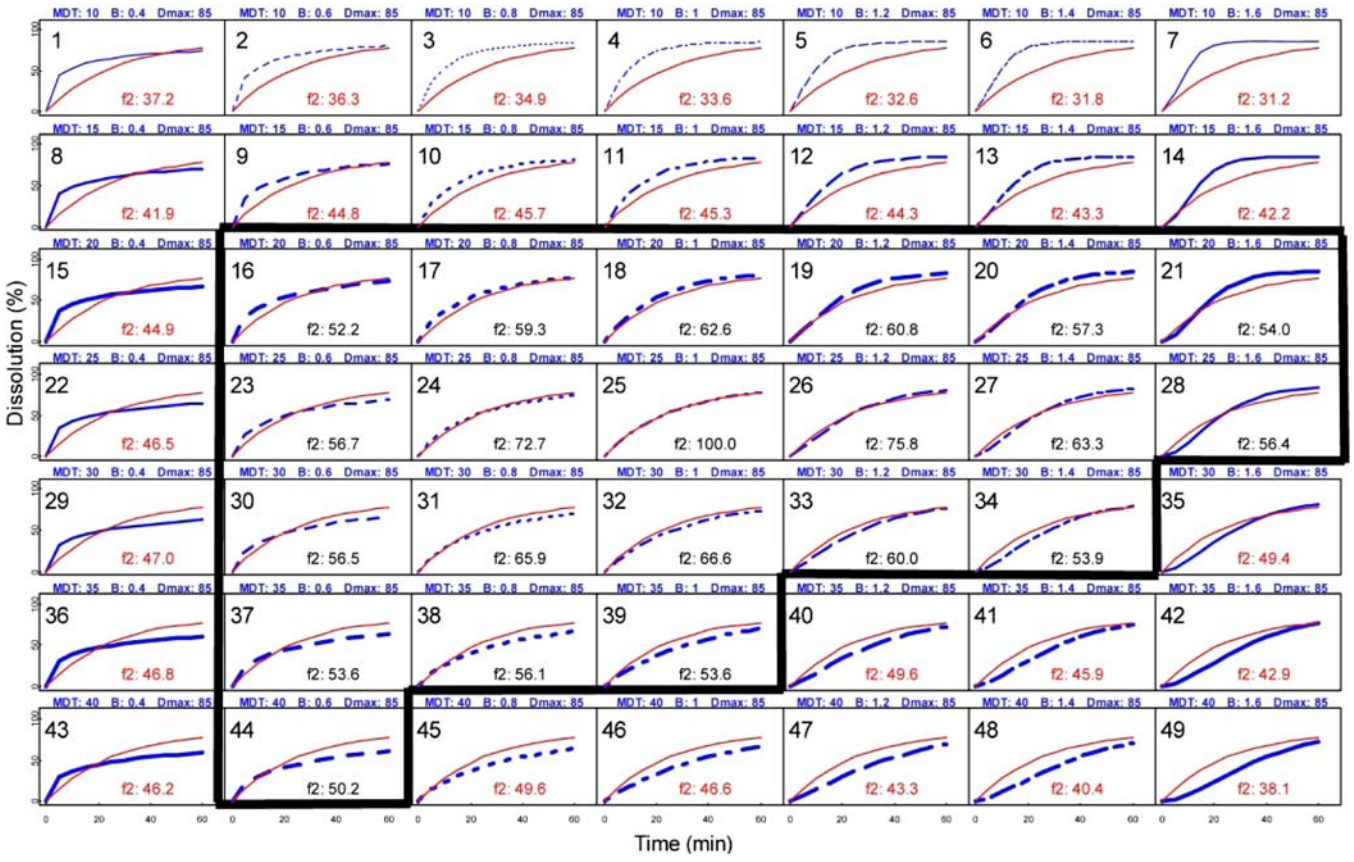


FIGURE 13.23 Simulated dissolution profiles. The parameters used for simulation are labeled at the top of each panel. The f2 similarity value is noted in each panel. f2 values ≥50 are colored black, while f2 values <50 are colored red (gray in print versions). The solid black border around the panels outlines the profiles that are similar.

product deviate from those of the reference, the profiles of the test product may be higher (top panels, such as panel 4), or lower than the reference (bottom panels, such as panel 46). In some other instances, the early parts of the test profile are higher, while the later parts are lower than the reference profile (the panels on the left, such as panel 22) or vice versa (the panels on the right, such as panel 28). The f2 values are labeled in each panel, which are smaller when the parameters deviate further. Those values that are less than 50 are colored red (gray in print versions), and thick black lines are drawn to separate the similar and dissimilar profiles.

For each dissolution profile, the in vivo concentration profile is predicted using an IVIVC model. The corresponding in vivo profiles for the dissolution profiles shown here are presented in Fig. 13.24. In each panel of this figure, the ratios of AUC and C_{max} (and the T_{max} difference) between the test and the reference are labeled, with those outside the BE range (0.8−1.25) colored red (gray in print versions). Thick black lines are drawn to separate the bioequivalent from biononequivalent profiles.

This study demonstrated the general consistency between in vitro dissolution comparison using the f2 factor and in vivo BE results. The results also indicate that dissolution profiles that are judged to be similar using the f2 factor may not always be bioequivalent when tested in vivo. On the other hand, in vitro dissolution profiles judged to be dissimilar by the f2 factor may sometimes generate in vivo bioequivalent profiles. As shown in Fig. 13.25, the Weibull parameter range used for generating the dissolution profiles, which are determined to be similar to the reference profile by f2 factor, is called the *f2 similarity region* (enclosed by thick, dashed blue (dark gray in print versions) lines), while the parameter range generating in vivo profiles that are bioequivalent to reference profiles is called the *bioequivalence region* (enclosed by red (light gray in print versions) solid lines for AUC, and black dotted lines for C_{max}). It can be seen that the f2 similarity region and the BE region are close, although they do not exactly match.

This study emphasized the importance of evaluating the shape and the completeness of in vitro dissolution curves when f2 is used to determine the similarity

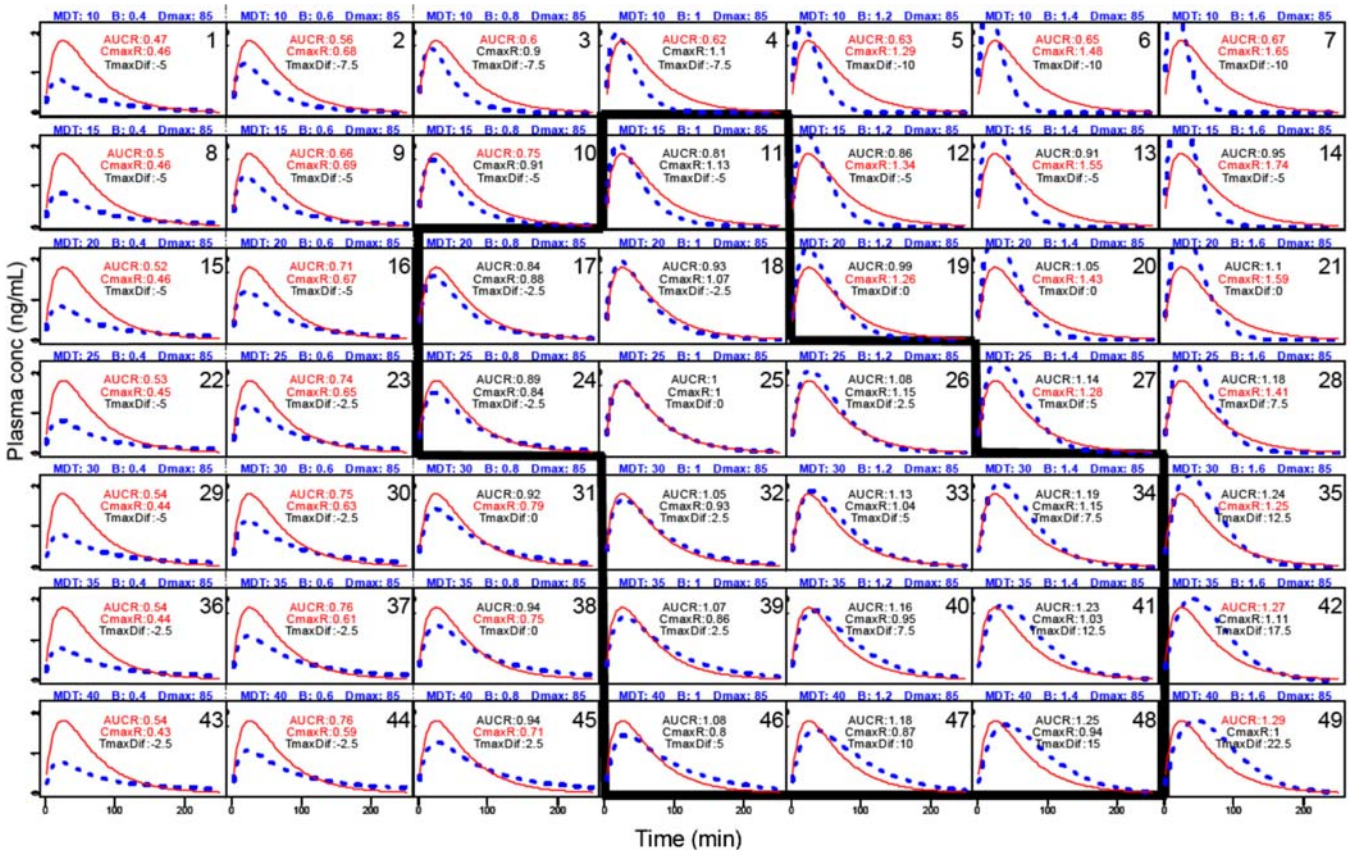


FIGURE 13.24 The in vivo profiles corresponding to the in vitro profiles in Fig. 13.22 based on the IVIVC model. At the top of each panel are the model parameters used for simulating the in vitro dissolution from which the in vivo profile were obtained. The red (gray in print versions) solid line denotes the reference profile, and the dotted line denotes the simulated profile. The AUC ratio, C_{max} ratio, and T_{max} difference are noted in each panel. The solid black border around the panels outlines the profiles that are bioequivalent.

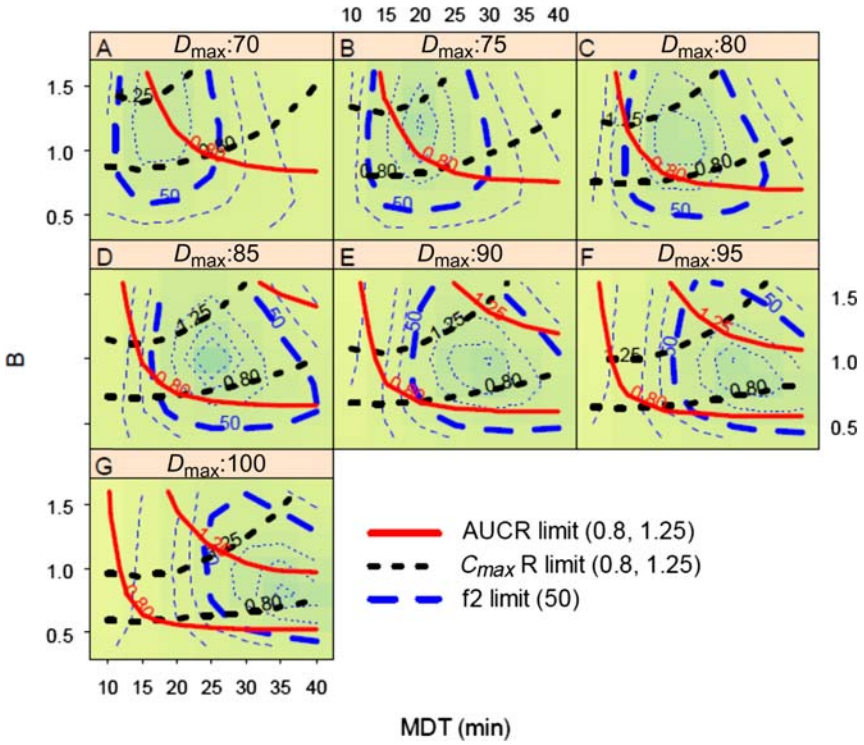


FIGURE 13.25 Effect of Weibull parameters (MDT, B, and D_{max}) used in the simulation of dissolution profiles on the AUC ratios, C_{max} ratios, and f2 comparisons. When the MDT and B take the values enclosed by the two thick red (light gray in print versions) lines (for AUCR) and the two black thick dotted lines labeled (for $C_{max}R$), the simulated in vivo concentration-time profiles are bioequivalent to the reference (BE region). The blue (dark gray in print versions) thick lines enclose a region in which the combinations of MDT, B, and D_{max} produce dissolution profiles similar to the reference profile (f2 similarity region).

between different formulations. This evaluation is important because the completeness of dissolution relates to the extent of drug absorption in vivo and the shape of a dissolution curve is translated to the rate of drug absorption in vivo. In particular, when there is a difference of more than 10% in the plateau levels between dissolution profiles of the test and reference product, or when the two dissolution profiles cross, there is a greater likelihood for the test product not to be bioequivalent to the reference product, although f2 similarity has been demonstrated. Under these circumstances, caution must be exercised when drawing conclusions.

13.5 SUMMARY

In this chapter, in vitro dissolution methods are developed to evaluate the in vivo bioperformance of solid oral dosage forms via IVIVCs, and as quality control tests to ensure the consistency of product manufacturing and the bioperformance of the resulting products. Valid IVIVCs are useful for biowaiver, for quality control, or both. Much research has been done since the middle of the 20th century, with successful IVIVCs being developed for some products. However, IVIVCs are still not yet established for many products, particularly for many ANDAs, either due to the fact that the complexity of drug dissolution and absorption along the GI tract is not yet fully understood, or the in vitro conditions could not mimic that of the GI system. Meaningful results and interpretation of dissolution data can be achieved only when the responses of the biological system to the physical properties of drug products are well understood, and test methods are properly established through thorough studies during the formulation and manufacturing process design and clinical development.

To provide meaningful dissolution data, numerous dissolution devices and methods (either compendial or not) have been studied for a variety of products, which will continue into the future with an attempt to provide biorelevant dissolution data. For any given apparatus, dissolution results can still be affected by various elements for the dissolution tests, such as setting and calibration of the apparatus, mixing speed, medium (volume, pH, ionic strength, surfactant, cosolvent, aeration), temperature, positions of dosage forms in the medium, sampling locations and time control, and interference by the environment. The effects of these elements and factors must be evaluated during the development of dissolution methods for each product.

The use of dissolution tests for quality control should also be directed toward guarding the bioperformance of products. For quality control purposes, the test methods should have sufficient discriminating ability to identify the changes in products and manufacturing processes that may lead to significant difference in the BA/equivalency than that of the pivotal batches or the BE batches. For this purpose, the effects of formulation and process parameters on product dissolution should be characterized so that discriminating methods can be developed to relate the dissolution behavior to changes in either formulation or process. However, the discriminating power of the methods should be evaluated for relevance to product bioperformance in order to make the methods and results meaningful. In the meantime, dissolution methods should be robust to allow normal variability in routine manufacturing and testing, and to avoid overdiscriminating, which may lead to unnecessary rejection of products.

Importantly, product dissolution could be highly method-dependent as well as product dependent. Some of the dissolution data may be correlated to the bioperformance of the product, while others may not. In the absence of IVIVCs, the quality meaning of dissolution data or similarity factor f2 cannot be defined for the bioperformance of the products. In fact, simulation studies for a product with valid IVIVCs have shown that not all batches of the product with $f2 > 50$ could be bioequivalent, and on the other hand, not all the batches with $f2 < 50$ fail to be bioequivalent. Therefore, unless the compositions and manufacturing processes are exactly the same, or unless IVIVCs have been established, the requirement in some countries for multisource products to have a similar dissolution before BE study is neither meaningful nor necessary, except for some rapid-release IR products of BCS class I and III drug substances, for which biowaivers may be possible if justified properly.

Specifications of product dissolution are always part of the dissolution tests. For NDAs, dissolution specifications are set based on the dissolution data of all clinical batches that have produced acceptable PK or clinical results. For ANDAs, specifications are set based on the dissolution data of the BE batches. Careful evaluation of the dissolution data of all the clinical/BE batches can provide an overall picture of the product behavior to facilitate the setting of specifications. ICH and FDA have provided guidance for setting dissolution specifications for different types of dosage forms, with or without IVIVCs. For products without IVIVCs, FDA guidance recommends setting the acceptable dissolution ranges as being $\pm 10\%$ of the clinical or bio-batches for all time points. When valid IVIVCs are available, a wider dissolution ranges can be set such that, via the IVIVC, the predicted C_{max} and AUC are in the range of $\pm 20\%$ of the target. For a

tighter control, the acceptance criteria is recommended to be set based on the average in vitro dissolution data equivalent to USP Stage 2 testing ($n = 12$) for each lot under study. Simulation is a useful tool to demonstrate the effects of mean and variability on the failure rate to pass different stages of dissolution test with desired power, which can also be used to examine if the proposed dissolution specifications are set properly.

References

1. USP38 Chapter <1088>.
2. FDA Guidance for Industry. Extended release oral solid dosage forms: development, evaluation, and application of in vitro/in vivo correlations; September 1997.
3. USP38 Chapter <1090>: Assessment of drug product performance-bioavailability, bioequivalence, and dissolution.
4. FDA Draft Guidance for Industry: Industry. Waiver of in vivo bioavailability and bioequivalence studies for immediate-release solid oral dosage forms based on a biopharmaceutics classification system; 2015.
5. FDA Guidance for Industry. SUPAC-IR: immediate release solid oral dosage forms scale-up and postapproval changes: chemistry, manufacturing, and controls, in vitro dissolution testing, and in vivo bioequivalence documentation; 1995.
6. FDA Guidance for Industry. SUPAC-MR: modified release solid oral dosage forms: scale-up and postapproval changes: chemistry, manufacturing, and control; in vitro dissolution testing, and in vivo bioequivalence documentation; September 1997.
7. FDA Guidance for Industry. FDA guidance for industry: dissolution testing of immediate release solid oral dosage forms; 1997.
8. EMA Draft Guideline. Quality of oral modified release products; 2012.
9. EMA Guideline. Investigation of bioequivalence; 2010.
10. WHO Working Document QAS/14.583. Multisource (generic) pharmaceutical products: guidelines on registration requirements to establish interchangeability; 2014.
11. MHLW Guideline for Bioequivalence Studies of Generic Drug Products. <www.jga.gr.jp/englsih/country-overview/bioequivalence-be>; February 2012.
12. Dokoumetzidis A, Macheras P. A century of dissolution research: from Noyes and Whitney to the biopharmaceutics classification system. *Int J Pharm* 2006;**321**:1–11.
13. Edwards LJ. The dissolution and diffusion of aspirin in aqueous media. *Trans Faraday Soc* 1951;**47**:1191–210.
14. Sinko PJ, Leesman GD, Amidon GL. Predicting fraction dose absorbed in humans using a macroscopic mass balance approach. *Pharm Res* 1991;**8**:979–88.
15. Oh DM, Curl RL, Amidon GL. Estimating the fraction dose absorbed from suspensions of poorly soluble compounds in humans: a mathematical model. *Pharm Res* 1993;**10**:264–70.
16. Amidon GL, Lennernas H, Shah VP, Crison JR. A theoretical basis for a biopharmaceutic drug classification: the correlation of in vitro, drug product dissolution and in vivo bioavailability. *Pharm Res* 1995;**12**:413–20.
17. Agrawal S, et al. Solid-state characterization of rifampicin samples and its biopharmaceutic relevance. *Eur J Pharm Sci* 2004;**22**:127–44.
18. Dabbagh MA, et al. Effects of polymer particle size, compaction pressure and hydrophilic polymers on drug release from matrices containing ethylcellulose. *Int J Pharm* 1996;**140**:85–95.
19. Kapsi SG, Ayres JW. Processing factors in development of solid solution formulation of itraconazole for enhancement of drug dissolution and bioavailability. *Int J Pharm* 2001;**229**:193–203.
20. Law D, et al. Ritonavir-PEG 8000 amorphous solid dispersions: in vitro and in vivo evaluations. *J Pharm Sci* 2004;**93**:563–70.
21. Draguet-Brughmans M, et al. The physicochemical properties of digoxin. *J Pharm Biomed Anal* 1985;**3**:227–34.
22. Landín M, et al. Influence of microcrystalline cellulose source and batch variation on the tabletting behaviour and stability of prednisone formulations. *Int J Pharm* 1993;**91**:143–9.
23. Wang LH, Chowhan GT. Drug-excipient interactions resulting from powder mixing. V. Role of sodium lauryl sulfate. *Int J Pharm* 1990;**60**:61–78.
24. Tobiska S, Kleinebudde P. Coating uniformity and coating efficiency in a Bohle Lab-Coater using oval tablets. *Eur J Pharm Biopharm* 2003;**56**:3–9.
25. Noyes AA, Whitney WR. The rate of solution of solid substances in their own solutions. *J Am Chem Soc* 1897;**19**:930.
26. Brunner E. Reaktionsgeschwindigkeit in heterogenen Systemen. *Z Phys Chem* 1904;**47**:56–102.
27. Nernst W. Theorie der Reaktionsgeschwindigkeit in heterogenen Systemen. *Z Phys Chem* 1904;**47**:52.
28. Yalkowsky SH, Valvani SC. Solubility and partitioning I: solubility of nonelectrolytes in water. *J Pharm Sci* 1980;**69**:912–22.
29. Soltero RA, et al. Effects of sinker shapes on dissolution1 profiles. *J Pharm Sci* 1989;**78**:35–8.
30. Kukura J, et al. Understanding pharmaceutical flows. *Pharm Technol* 2002;**26**:48–72.
31. Kukura J, et al. Engineering tools for understanding the hydrodynamics of dissolution tests. *Drug Dev Ind Pharm* 2003;**29**:231–9.
32. Gao Z, et al. Studies of variability in dissolution testing with USP apparatus 2. *J Pharm Sci* 2007;**96**:1794–801.
33. Healy AM, et al. Sensitivity of dissolution rate to location in the paddle dissolution apparatus. *J Pharm Pharmacol* 2002;**54**:441–4.
34. Tang L-J, Schwartz JB. Introduction of dissolution error as a result of different openings in vessel covers. *Pharm Dev Tech* 1998;**3**:261–7.
35. Missaghi S, Fassihi R. Release characterization of dimenhydrinate from an eroding and swelling matrix: selection of appropriate dissolution apparatus. *Int J Pharm* 2005;**293**:35–42.
36. Gao ZM. Mathematical modeling of variables involved in dissolution testing. *J Pharm Sci* 2011;**100**:4934–42.
37. ASTM International. E2503-07; 2007.
38. Gao ZM, et al. Effects of deaeration methods on dissolution testing in aqueous media: a study using a total dissolved gas pressure meter. *J Pharm Sci* 2006;**95**:1606–13.
39. Gao ZM, et al. Effects of vessel geometric irregularity on dissolution test results. *J Pharm Sci* 2011;**100**:1093–101.
40. Gao ZM, et al. Vibration effects on dissolution tests with USP apparatuses 1 and 2. *J Pharm Sci* 2008;**97**:3335–43.
41. Gao ZM, et al. The random vibration effects on dissolution testing with USP apparatus 2. *J Pharm Sci* 2009;**98**:297–306.
42. Gao ZM, et al. Vibration effects of lab equipment on dissolution testing with USP paddle method. *J Pharm Sci* 2010;**99**:403–12.
43. Brown CK, et al. FIP position paper on qualification of paddle and basket dissolution apparatus. *AAPS PharmSciTech* 2009;**10**:924–7.
44. USP 35, Chapter <711>.
45. Foster T, Brown W. USP dissolution calibrators: re-examination and appraisal. *Dissolution Technol* 2005;**1**:6–8.
46. Subcommittee on Dissolution Calibration, Pharmaceutical Research and Manufacturers of America (PhRMA). Dissolution calibration: recommendations for reduced chemical testing and enhanced mechanical calibration. Pharm Forum 2000;**26**(4):1149–66.

47. Cecil T. Dissolution calibration and testing workshop. *Dissolution Technol* 1995;**2**(3):8.
48. FDA Guidance for Industry (2010). The use of mechanical calibration of dissolution apparatus 1 and 2—current good manufacturing practice (CGMP).
49. ASTM International. ASTM E2503-07; 2007.
50. European Pharmacopoeia (Ph. Eur.). Dissolution test. 6th ed. January 2008: 20903.
51. Japanese Pharmacopoeia (JP) XV, Chapter <6.10> Dissolution test.
52. ICH. Topic Q4B Annex 7 Dissolution test general chapter, step 3; 2008.
53. Iarriccio A, Kassis A, Patel T. Increasing the efficiency of the dissolution laboratory through automation. *Am Pharm Rev* 2011;**14**(4):1–10.
54. Kretz J, Wong–Moon K. Evaluation of automation to increase efficiency in the dissolution lab. *Dissolution Technol* 2013;**20**(2):33–7.
55. USP 35-NF 30, Chapter <1092>.
56. McAllister M. Dynamic dissolution: a step closer to predictive dissolution testing? *Mol Pharm* 2010;**7**(5):1374–87.
57. Mann JC, Pygall SR. A formulation case study comparing the dynamic gastric model with conventional dissolution methods. *Dissolution Technol* 2012;**19**(4):14–19.
58. Blanquet S, et al. A dynamic artificial gastrointestinal system for studying the behavior of orally administered drug dosage forms under various physiological conditions. *Pharm Res* 2004;**21**(4):585–91.
59. Mercuri A, et al. Assessing drug release and dissolution in the stomach by means of dynamic gastric model: a biorelevant approach. *J Pharm Pharmacol* 2009;**61**(S1):A5.
60. Minekus M, Havenaar R. US Patent 5525305; 1996.
61. Rich GT, et al. Solubilization of carotenoids from carrot juice and spinach in lipid phases: I. Modeling the gastric lumen. *Lipids* 2003;**38**:933–45.
62. Mercuri A, et al. Dynamic gastric model (DGM): a novel in vitro apparatus to assess the impact of gastric digestion on the droplet size of self-emulsifying drug-delivery system. *J Pharm Pharmacol* 2008;**60**(S1):A2.
63. Buch P, et al. IVIVC in oral absorption for fenofibrate immediate release tablets using a dissolution/permeation system. *J Pharm Sci* 2009;**98**(6):2001–9.
64. Yano K, et al. Mechanisms of membrane transport of poorly soluble drugs:role of micelles in oral absorption processes. *J Pharm Sci* 2010;**99**(3):1336–45.
65. Ginski MJ, et al. Prediction of dissolution-absorption relationships from a continuous dissolution/Caco-2 system. *AAPS PharmSci* 1999;**1**:E3.
66. Gao Z. Development of a continuous dissolution/absorption system—a technical note. *AAPS PharmSciTech* 2012;**13**(4):1287–92.
67. Selen A, et al. *Application of FloVitro™ technology to evaluate dissolution of furosemide and danazol in simulated media at fed and fasted conditions*. Washington DC: AAPS; 2011. p. 10.
68. Garbacz G, et al. Irregular absorption profiles observed from diclofenac extended release tablets can be predicted using a dissolution test apparatus that mimics in vivo physical stresses. *Eur J Pharm Biopharm* 2008;**70**(2):421–8.
69. Gao Z. *In-vitro dissolution testing of gelatin capsules with applied mechanical compression*, 10. Orlando, FL: AAPS; 2015.
70. ICH Harmonized Tripartite Guideline. Specifications: Test Procedures and Acceptance Criteria for New Drug Substances and New Drug Products: Chemical Substances Q6A.
71. ICH Harmonized Tripartite Guideline. Q8: Pharmaceutical Development.
72. FDA Draft Guidance for Industry. Bioavailability and bioequivalence studies submitted in NDAs or INDs—general considerations; 2014.
73. Code of Federal Regulations (CFR), Section 21CFR320.24.
74. FDA Draft Guidance for Industry. Bioavailability and bioequivalence studies submitted in NDAs or INDs—general considerations; 2014.
75. Digenis GA, Gold TB, Shah VP. Cross-linking of gelatin capsules and its relevance to their in vitro-in vivo performance. *J Pharm Sci* 1994;**83**(7):915–21.
76. FDA. Dissolution methods. <http://www.accessdata.fda.gov/scripts/cder/dissolution>; 2016 [accessed 7/8/2016].
77. Anand O, et al. Dissolution testing for generic drugs: an FDA perspective. *AAPS J* 2011;**13**(3):328–35.
78. FDA Draft Guidance for Industry. Bioequivalence studies with pharmacokinetic endpoints for drugs submitted under an ANDA; 2013.
79. Moore J, Flanner H. Mathematical comparison of dissolution profiles. *Pharm Technol* 1996;**20**:64–75.
80. O'Hara T, et al. Review of methodologies for the comparison of dissolution profile data. *Adv Exp Med Biol* 1997;**423**:167–71.
81. Polli J, et al. Methods to compare dissolution profiles and a rationale for wide dissolution specifications for metoprolol tartrate tablets. *J Pharm Sci* 1997;**86**:690–700.
82. Costa P, Lobo JS. Modeling and comparison of dissolution profiles. *Eur J Pharm Sci* 2001;**13**:123–33.
83. Shah V, et al. In vitro dissolution profile comparison—statistics and analysis of the similarity factor, f2. *Pharm Res* 1998;**15**:889–96.
84. Duan JZ, Riviere K, Marroum P. In vivo bioequivalence and in vitro similarity factor (f2) for dissolution profile comparisons of extended release formulations: how and when do they match? *Pharm Res* 2011;**28**(5):1144–56.

CHAPTER

14

Bioavailability and Bioequivalence*

H. Zhu, B.V. Li, R.S. Uppoor, M. Mehta and L.X. Yu

U.S. Food and Drug Administration, Silver Spring, MD, United States

14.1 GENERAL BACKGROUND

Bioavailability (BA) assessment is essential in oral dosage form development.[1] BA studies, which characterize overall rate and extent of drug absorption, are commonly conducted to support new drug applications (NDAs). Different types of BA studies are carried out for various reasons:

- A relative BA study may be required to assess the exposure similarity between a new oral dosage form and an approved product (ie, reference list drug) in a development program for a 505 b(2) product. The exposure difference between the two products is useful to select doses in a subsequent clinical trial aimed to obtain efficacy and safety information about the new product.
- Demonstration of sufficient systemic absorption is essential to justify the necessity of formulating a new compound, such as a new chemical entity, into an oral dosage form. A dedicated BA study may be necessary by comparing systemic exposure of the new oral dosage form to that of a reference formulation, such as an intravenous formulation. BA information can also be obtained indirectly through multiple early phase clinical trials, such as a single ascending dose study, a multiple ascending dose study, or a mass balance study.
- Sometimes BA can be altered due to food intake or concomitant medications. The effect of food on absorption is typically investigated in a dedicated study in order to provide dosing instructions.[2]
- Likewise, for a drug product whose dissolution is significantly affected by pH (eg, over the pH range from 1 to 6.5), a drug–drug interaction study with a pH altering agent, such as a proton pump inhibitor, may be appropriate to evaluate the potential changes on BA.[3]

Bioequivalence (BE) studies focus on demonstration of absence of difference in the rate and extent of two products.[1,4] In other words, two bioequivalent products are anticipated to be similar in BA. BE studies can be used to support approvals of NDAs, abbreviated NDAs (ANDAs), and postapproval changes, illustrated as following:

- BE studies can be applied to support NDA submissions under several situations. One common situation is that data from a BE study can be applied to bridge a clinical trial formulation and a to-be-marketed formulation whose features are improved in parallel with the clinical development. In another situation, a new product, which is developed as a new oral dosage form or uses a new salt form of an approved active moiety, can gain marketing access relying on the agency's efficacy and safety findings of an approved product if BE between the new product and the existing product can be established.

*__Disclaimer__: The chapter reflects the views of the authors and should not be construed to represent the views or policies of the Food and Drug Administration (FDA). No official endorsement by the FDA is intended or should be inferred.

DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00014-5

- Generic substitutes, which are therapeutically equivalent and often less expensive to brand name products, are approved under ANDAs. BE must be demonstrated to support approvals of ANDAs.
- BE studies may also be required during postapproval period of NDAs or ANDAs in certain situations, for example when a major change in formulation, or manufacturing process, or manufacturing site occurs.

The concept of BA and BE has been evolving over years. In the 1970s, it was reported that patients receiving digoxin showed large variation in clinical responses. Subsequent investigations indicated that different clinical outcomes were associated with variation in BA of products manufactured by different firms or by the same firm but from different lots.[5,6] In 1974, one committee organized by the FDA Office of Technology Assessment (OTA) concluded that "current standards and regulatory practices do not assure bioequivalence."[7,8] In 1977, the Food and Drug Administration (FDA) started to request in vivo assessment in its published regulation on BE.[9] In 1984, the US Congress passed the Hatch-Waxman Act (also known as "Drug Price Competition and Patent Term Restoration Act"), which provides a legal basis to allow a BE study as a surrogate for clinical efficacy and safety trials. Since then, FDA published BE guidances for various specific products in addition to its general BA/BE guidance. BA/BE studies have become standard assessment in clinical development programs and support drug product approvals.[10] In recent years, BA/BE assessment has been expanded to address drug products with unique complexities, such as for drugs with narrow therapeutic indexes (NTIs), products with multimodal release mechanisms, and products acting locally.

The information included in this chapter will provide readers with an overview of the general concept of BA and BE. Details on typical BA/BE study designs and data analyses will be discussed, with a primary focus on the orally administered drugs. Special topics on BE for NTI drugs and highly variable (HV) drugs, together with application of partial area under the curves (pAUCs) for BE testing will also be presented.

14.2 DEFINITIONS AND KEY CONCEPTS

Different oral formulations may be associated with different absorption rate and extent. An immediate release formulation is designed to allow rapid release of an active moiety. An extended release formulation, on the other hand, controls drug release rate to ensure active moiety is available over a prolonged time interval. In recent years, a more complicated formulation is designed to include both an immediate-release component and an extended-release component. Different formulations intended for different BA features usually yield different plasma concentration time profiles.

BA/BE assessment is established on the basis of pharmacokinetics (PK), which characterizes plasma drug concentration versus time profiles. PK provides information on ADME process (ie, absorption, distribution, metabolism, and excretion) and describes the fate of an administered drug. Key PK parameters such as C_{max}, T_{max}, and area under the curve (AUC) can be obtained through concentration−time profiles. C_{max} represents the maximum concentration of the curve. The time to reach C_{max} is known as T_{max}. If adequately frequent plasma samples are taken in an appropriately designed study, C_{max} and T_{max} can be directly observed. AUC is the area under the concentration-time curve for the time interval of interest and is calculated by applying a numerical integration formula such as the trapezoidal rule. AUC_{0-t} represents the AUC over the time interval from zero to the last observation, while $AUC_{0-\infty}$ represents the AUC over the time interval from zero to infinity. $AUC_{0-\tau}$ is the AUC over a dosing interval (ie, τ). In the following section, the key concepts of BA and BE and the assessment of BA and BE using the major PK parameters are provided.[11]

14.2.1 Bioavailability

As defined in the Code of Federal Regulations, Title 21, Section 320, Part1 (§ 320.1), BA is known as "the rate and extent to which the active ingredient or active moiety is absorbed from a drug product and becomes available at the site of action."[12] Most drugs administered orally are not intended to be used locally in the gastrointestinal (GI) tract; instead, pharmacological effects can manifest only after the drug is absorbed and an adequate amount becomes available at the intended receptors of the target tissue through systemic circulation. Direct assessment of drug concentration at the target tissue is technically challenging. In addition, drug concentration at the local action site is thought to be in equilibrium with the drug concentration in systemic circulation. Therefore, in practice, the drug concentration in systemic circulation is used as a surrogate to that of the drug presented at the target tissue.

For an oral product, BA is mainly affected by three factors: physicochemical characteristics of a drug, dosage factors, and physiological factors (eg, pH in the GI tract).[13,14] Different studies can be conducted to evaluate changes in BA, which is commonly described by a ratio of C_{max} or AUC, due to these factors.

A BA study of an oral product estimates the relative fraction of the administered dose that is absorbed into systemic circulation when compared to that of a reference formulation. Ideally, a reference formulation is an optimally available formulation, which can be an intravenous formulation, an oral solution, or an oral suspension. Sometimes a different oral dosage form (eg, a different oral tablet) can be used too. Absolute BA is used to describe the relative fraction of systemic exposure for an oral dosage form when an intravenous dosage form (100% bioavailable) is used as the reference formulation, whereas the term *relative BA* is used when the reference is another nonintravenous formulation.[1]

PK parameter ratios, such as AUC ratios, between two formulations are commonly used to compare BA difference between two formulations. $AUC_{0-\infty}$ (or AUC_{0-t}) is generally applied for a single-dose BA study, whereas $AUC_{0-\tau}$ is for a multiple-dose study. Eq. (14.1) is used to calculate absolute BA, where D_{iv} and D_{po} are the intravenous and oral doses administered, respectively; and AUC_{iv} and AUC_{po} are the AUC estimates for the intravenous and oral routes, respectively.[1]

$$F = \frac{D_{iv}}{D_{po}} \cdot \frac{AUC_{po}}{AUC_{iv}} \tag{14.1}$$

Relative BA can be calculated by using Eq. (14.2), where D_A and D_B are the doses administered for drug formulation A and B, respectively; and AUC_A and AUC_B are the AUC estimates for the A and B formulations, respectively. Formulation A is the reference formulation:

$$F_{rel} = \frac{D_A}{D_B} \cdot \frac{AUC_B}{AUC_A} \tag{14.2}$$

14.2.2 Bioequivalence

According to § 320.1 (21CFR320.1), BE is defined as "the absence of a significant difference in the rate and extent to which the active ingredient or active moiety in pharmaceutical equivalents or pharmaceutical alternatives becomes available at the site of drug action when administered at the same molar dose under similar conditions in an appropriately designed study."[12] Similar to BA assessment, BE assessment for oral dosage forms can be performed based on drug concentration in systemic circulation, which is considered as a surrogate for concentration available at the site of action. Drug concentration in systemic circulation is determined by the ADME process, in which only absorption is thought to be affected by a formulation. Similar concentration levels are anticipated in the same subject receiving different oral dosage forms that show similar absorption rate and extent.

C_{max} and AUC, which are applied for BA evaluation, are the two major PK parameters used for BE assessment.[1] C_{max} bears the information on absorption rate, where an increased absorption rate is associated with an elevated C_{max}. The similarity in C_{max} of two formulations is suggestive for similar maximum effects or similar risks of the safety events. Like BA evaluation, $AUC_{0-\infty}$ (or AUC_{0-t}) and $AUC_{0-\tau}$are used to describe overall exposure for a single-dose BE study and a multiple-dose BE study, respectively. Generally, AUC represents the extent of absorption and is independent of the absorption rate. For drugs requiring chronic treatment, such as antidepressants, AUC is correlated with effectiveness and long-term safety events. Two products yielding similar AUCs are thought to have similar efficacy and long-term safety profiles. Other PK parameters, such as pAUC and T_{max}, are also used for BE assessment.

Pharmacodynamic, clinical, or in vitro assessment may also be performed in BE assessment.[4,15,16] Under situations where systemic absorption is not directly indicative of drug available at its site of action, a pharmacodynamic, clinical, or in vitro assessment may be used. This approach is commonly seen in small molecule products for local actions, such as topical products, nasal and inhalation products, locally acting GI products, and dental products. It is rarely seen in oral dosage forms intended for systemic indications.

14.2.3 Pharmaceutical equivalents, pharmaceutical alternatives, and therapeutic equivalents

Drug products are considered to be pharmaceutical equivalents if they contain the same active ingredient, in the same amount, with identical dosage forms and routes of administration. Pharmaceutical alternatives are drug products containing the same therapeutic moiety, but are different salts, esters, or complexes of that moiety, or are different dosage forms or strengths. Furthermore, drug products are considered to be therapeutically equivalent only if they are pharmaceutical equivalents (as described previously) that are expected to produce the same clinical effects and have similar safety profiles when they are administered to patients under the same conditions as specified in the product labeling information. In general, therapeutic equivalence may be reasonably inferred if two pharmaceutical equivalent products are shown to be bioequivalent. Therapeutic equivalence is thus an ultimate measure of the interchangeability of two distinct drug products or formulations.[10]

14.3 GENERAL COMPONENTS OF BA AND BE STUDIES

In clinical development, demonstration of BE and assessment of BA for oral dosage forms intended for systemic indications generally require dedicated phase 1 clinical trials using PK parameters as the endpoints. The focus for BA or BE studies is to evaluate absorption rate and extent, even though the study objectives differ. Therefore, design features are similar between the two types of clinical trials. In the following sections, the design features of a BE study will be discussed, as an example, in detail.

14.3.1 Study population

Standard BE studies are conducted in adults at least 18 years old and capable of giving informed consent. For most drugs, the PK processes are affected by demographic characteristics of patients. Thus, demographic characteristics of subjects enrolled in a BE study should represent the patient population, taking into account age, sex, and race. For example, if an oral dosage form is intended for both male and female patients, a BE study should be conducted in subjects with both sexes.[1,4]

Healthy volunteers are generally preferred in a BE study over patients for several reasons. BE assessment focuses on a comparison of drug absorption rate and extent of two formulations, not the therapeutic effectiveness of a formulation. The physical processes of drug absorption for solid oral dosage forms are usually the same in patients as they are in healthy subjects. Therefore, healthy subjects can adequately identify the differences in formulations, which fulfills the BE study purpose. On the other hand, patients may have different severity levels of the diseases, and some may require concomitant medication treatment, all these factors may potentially confound the BE study outcome. Practically, healthy volunteers are easy to enroll with simple inclusion and exclusion criteria and less chance of using concomitant medications. In addition, intensive PK samples are usually better tolerated in healthy volunteers than patients.

Sometimes patients are needed for a BE study. Safety and tolerability risks associated with the drug product are the general reasons that preclude healthy volunteers being enrolled in a BE study. For example, clozapine is approved for treatment-resistant schizophrenia. Clozapine has been shown to cause several potentially fatal adverse events, including agranulocytosis, syncope, seizures, myocarditis, and cardiomyopathy.[17] Because of these safety risks, FDA guidance recommends that the BE studies be conducted in stable patients.[18]

14.3.2 Study design

One typical BE study design is a single-dose, two-way crossover study design.[1,4,19] As illustrated in Table 14.1, subjects enrolled in the study are randomized into two sequence groups (Groups TR and RT) with balanced baseline and demographic characteristics. After the trial is initiated, subjects in the two groups receive test products and reference products, respectively. A reference product is usually the product whose efficacy and safety profile has been established, whereas a test product is the new product. If the drug is well tolerated, a subject is usually given a single unit of the oral dosage form with the highest strength. Blood samples are collected in subjects from both groups at various time points for PK characterization. Usually, the PK sampling should continue for at least three half-lives of the drug. To minimize potential carryover effect, time for wash-out accounting for at least five half-lives of the drug is recommended before the second period of dosing is initiated. In the second period, the treatment is switched. Subjects receiving the test product in period 1 are given the reference product in period 2, and vice versa. Blood samples are collected in the same manner as in period 1.

One feature of a crossover study is the same individual receives both the test product and the reference product, and the formulation effect can be compared at an individual level. It has been shown that the ADME process varies among subjects and may affect the variability of C_{max} and AUC estimates. The same subject provides a relatively homogenous environment to assess the formulation effect. Statistically, this approach is powerful and requires small sample size.

For a product with long half-life, a parallel study is practically more appealing than a crossover study.[1,4,19] Subjects enrolled in a parallel study are generally randomized into two groups with test product and reference product administered to one of the groups, respectively. Blood samples are collected for PK assessment over the time interval of at least three half-lives of the drug. In general, a parallel study requires less time than a crossover study because a subject receives treatment once and no washout is necessary. However, a large sample size is needed in a parallel

TABLE 14.1 Diagram of a Typical Two-Treatment, Two-Way Crossover BE Study

	Sequence group	
Period	**1**	**2**
1	T	R
2	R	T

Note: T, test product; R, reference product.

study to identify the formulation effect because the difference in ADME process across different individuals may increase the variability of the estimated C_{max} and AUC.

Through simulations and clinical trials, it has been shown that a single-dose BE study is generally more sensitive than a multiple-dose study. For drugs with a high accumulation factor, a smaller confidence interval was shown at steady state than that following a single dose. Therefore, there is a better chance to claim BE based on results from a steady-state study than from a single-dose study. This is unfavorable from the regulatory standpoint of view.[20] However, for practicality reasons, a steady-state BE study may still be required. Usually, a steady-state BE study is conducted in patients stabilized with the treatment. It is not feasible to discontinue the current treatment or to completely wash out the patients. Take clozapine, for example. Per the FDA guidance, the BE study should be conducted in patients receiving multiples of 100 mg of an approved clozapine twice daily. Once entered the study (ie, in period 1), the patients should be dosed either with the test product or the reference product twice daily for 10 days before blood sample collection. Then the patients switch the treatment for another 10 days in the second period.[18]

Sometimes a replicated crossover study may be conducted to assess intrasubject (within-subject) variability allowing a scaled BE approach.[1,4,19] A typical replicate BE study is illustrated in Table 14.2, where subjects enrolled in the study may receive the same treatment more than once. As shown in Table 14.2a, subjects in group 1 receive the test product twice and subjects in group 2 receive the reference product twice. This study is known as *partial replication*. A partial replicate design with two treatments on reference product is commonly used in BE studies for HV drugs. Table 14.2b shows that subjects in two treatment groups receive both the test product and the reference product twice but with different order. This study is known as *full replication*. A full replicate design is preferred in a BE study for an NTI drug.

TABLE 14 2 Diagram of a Typical Replicate BE Study Design

(a) Partial replication

	Period		
Sequence	**1**	**2**	**3**
1	T	R	T
2	R	T	R

(b) Full replication

	Period			
Sequence	**1**	**2**	**3**	**4**
1	T	R	T	R
2	R	T	R	T

Note: T, test product; R, reference product.

14.3.3 Biofluid matrices

PK endpoints used in BA/BE studies are commonly derived from concentration levels in peripheral blood (eg, plasma, serum, or whole blood). Blood samples are usually collected at defined time points to fully characterize a PK profile. Blood plasma, which contains extracellular fluid and protein, is the pale-yellow liquid after the collected blood samples are mixed with anticoagulant and blood cells are removed through centrifugation. Serum is the yellowish liquid after the whole blood is clotted. The content in serum is similar to plasma, except serum contains no clotting factors. Because most clotting factors do not specifically bind drugs, serum concentration and plasma concentration are usually similar. Drug concentration in whole blood includes drug concentration from both plasma and blood cells. Some drugs specifically bind to red blood cells. So the drug concentration in whole blood can be substantially different from that in serum or plasma.

Urine samples are sometimes taken for BA assessment. Urine samples can be useful if the drug is mainly eliminated through urine as the parent compound and the amount of drug excreted can be collected in a reasonably short period and measured as accurately as possible. It is uncommon to use other biofluids, such as saliva, cerebral spinal fluid, and lymphatic fluid in a BA/BE study.

14.3.4 Bioanalytical methods

In general, samples collected in a BA/BE study are quantified using a validated chromatography or a ligand binding assay.

High-performance liquid chromatography (HPLC) is the most commonly used chromatography assay to analyze PK samples. Usually, PK samples collected at defined time points are purified first. The objective for sample purification is to remove proteins, lipids, fatty acids, and other endogenous components that may affect sample separation and detection.[21] Proteins in the collected samples are precipitated by mixing with miscible organic solvents (eg, methanol or acetonitrile), often adjusted by different buffer, acid, or base, and removed through centrifugation. If needed, samples are further purified through liquid−liquid extraction (LLE)[21] and solid phase extraction (SPE).[22] For LLE,

immiscible organic solvents (eg, diethyl ether, MTBE) are mixed with the biological samples. The analyte of interest, usually transferred into the organic layer, is cleaned up for assay. For SPE, analyte is absorbed onto a solid phase column and then extracted by a diluent. Even with all cleanup procedures, the samples are still mixtures of various compounds. Second, the prepared samples are injected into chromatography instrument. Chromatography is a technique that the compound of interest is separated from interfering components, such as impurities, metabolites, and endogenous components remaining in the samples. The separation occurs from the interaction between the analyte and the analytical matrix (ie, a stationary phase and a mobile phase). In general, the separation relies on differences of hydrophobicity (reverse phase), molecular charge (ion exchange), and size (size exclusion) of the analyte and interfering components. Scientists may select appropriate solid phase and adjust mobile phase to ensure an optimal separation of the analyte and other components.[23] Third, the analyte is detected and quantified through a detection system. Various detection systems can be used in combination with a HPLC, where mass spectrometry is probably the mostly commonly used. Analytes obtained through a chromatography system are ionized first and then detected based on mass-to-charge ratio (m/z). Eventually, the concentration of the analyte can be obtained from chromatographic measurement.

A ligand assay relies on immunological reaction between an analyte and its specific antibody. In most cases, a ligand assay is applied to obtain concentrations of macromolecules, such as enzymes, peptides, and monoclonal antibodies.[24] Enzyme-linked immunosorbent assay (ELISA) is a good example. For a noncompetitive ELISA, the surface of a microplate is coated by a capture antibody to block all nonspecific binding sites. When the PK samples are applied, the analyte of interest is immobilized onto the plate in a concentration-dependent manner. The rest of the sample is then removed and the plate is washed. A secondary antibody containing an enzyme is added to the plate and binds to the analyte of interest, and then a specific enzyme substrate is added. The concentration of the analyte can thus be detected through the formation of the product of the enzyme and substrate interaction. The detection mechanism for a competitive ELISA is similar to that for a noncompetitive ELISA. The main difference is that a specific antigen is immobilized onto a microplate first. When an antibody is added, both antigens presented onto the plate and in the sample (ie, analyte of interest) compete for the limited binding sites of the antibody. Then, an antibody-analyte conjugate is washed away. Thus, a high concentration of analyte (antigen presented in the sample/solution) yield low antibodies immobilized onto the plate. The analyte concentration can thus be calculated.

Bioanalytical methods used in BA and BE studies, no matter a chromatography assay, a ligand assay, or other assays, are expected be accurate, precise, selective, sensitive, reproducible, and stability-indicating.[25]

Accuracy describes the closeness of the test results to the true concentration of the analyte. Accuracy is determined by replicated analysis of samples with known amount of the analyte using a specific analytical method. In general, a minimum of five determinations per concentration is recommended.

Precision describes the closeness of the test results upon multiple measurements with the same concentration level. There are two types of precision. One is known as within-run precision, which is an assessment of the precision during a single analytical run. The other one is known as *between-run precision*, which is a measurement of the precision with time, and may involve different analysts, equipment, reagents and laboratories. In general, the determined precision should not exceed 15% of the coefficient of variation (CV). At the lower limit of quantification, precision should not exceed 20% of the CV.

Selectivity is defined as the ability of an analytical method to differentiate and quantify an analyte in the presence of other components in the sample. Other compounds presented in the biological samples, including endogenous moieties, impurities, metabolites, and degraded products, may interfere with the analyte of interest. A validated method should demonstrate its ability to selectively quantify only the analyte of interest.

The stability of the chemical in certain matrix under specific conditions and in a specific time interval should also be assessed.

In 2013, FDA revised its guidance on bioanalytical method validation for bioanalytical methods mainly used in clinical pharmacology, BA, and BE studies. In this guidance document, the recommended validation information is discussed in detail.[25]

14.3.5 Compounds for bioassay

Concentrations of a parent compound (active ingredient or its active moiety in the dosage form) and its metabolites can be obtained through a validated bioassay.[1,4] The FDA recommends both the parent compound and, if possible, its active metabolites be measured in a BA study with the objective to assess the release of drug substance from a dosage form, and presystemic/systemic actions on the released drug substance. Concentration of a parent compound is

more preferable than metabolites in a BE study because changes in concentrations of a parent compound directly reflect the release rate and release extent of a dosage form. However, metabolite concentration may be preferred when the parent compound (eg, a prodrug) concentration is too low or measurable concentration lasts for too short a time. Sometimes both the metabolite and the parent compound concentrations are measured in a BE study when (1) the metabolite is shown to be formed in the GI tract or presystemically, and (2) both the parent compound and the metabolite contribute meaningfully to the overall efficacy or safety.

Concentrations of individual enantiomers can be obtained through a chiral assay. In general, FDA recommends that enantiomers be measured in a BA study.[26] BE assessment is usually performed based on racemate.[1] Sometimes enantiomers are measured if they exhibit different PK or pharmacodynamics characteristics, especially if the primary efficacy and safety reside in a minor enantiomer or nonlinear absorption is shown in one of the enantiomers.[1,27]

14.4 DATA ANALYSIS FOR BA AND BE STUDIES

14.4.1 Variables for BA/BE assessment

C_{max} and AUC are the most commonly used PK variables for in vivo BA/BE assessment.[1] As discussed in Section 14.2, C_{max} represents the rate of absorption and AUC represents the extent of absorption. Sometimes T_{max} and the shape of the PK profiles are also compared in a BA or BE assessment through visual inspection.

Quantitative analyses for the PK variables, such as C_{max} and AUC, are generally based on logarithmic transformation, which provides several advantages.[19] First, a direct subtraction of a logarithmic-transformed PK variable obtained from a test product and a reference product can be easily converted to the ratio of the same variable under normal scale. Usually, the ratio of the PK variable provides the most clinically relevant information. Second, PK variables can be expressed in a multiplicative fashion. Let's take $AUC_{0-\infty}$ as an example. It can be described by using Eq. (14.3), where D is the dose given to the subject. F and CL are the bioavailability and clearance for this individual, respectively:

$$AUC_{0-\infty} = \frac{F \cdot D}{CL} \tag{14.3}$$

The multiplicative terms (ie, F and CL) can thus be considered as a function of the subject. In the statistical analysis models for BE testing, all factors (eg, treatment, subjects, and sequence) that contribute to the variation in PK measurements are considered to be additive effects. Logarithmic transformation of AUC yields an additive equation (Eq. (14.4)):

$$\log AUC_{0-\infty} = \log D + \log F - \log CL \tag{14.4}$$

which is therefore preferable in BE analyses. Similar arguments can also be applied to C_{max} or other PK variables. Third, the observed AUC (or C_{max}) values in BA or BE studies under normal scale are often skewed and the variances between the test and reference groups may differ. The statistical models for BE testing are established based on two major assumptions: the normality and homoscedasticity of the variance. Log-transformation of the PK variables, however, makes the distribution appear more symmetric, closer to the normal distribution, and achieves a relatively homogeneous variance. The underlying assumptions for statistical testing can be better satisfied. Thus, log-transformations of PK variables in BA/BE evaluation is recommended by major regulatory bodies world-wide. It should be noted that log-transformation compares the geometric means of the PK variables from different treatment groups instead of arithmetic means.

Variability for the major PK parameters (eg, log AUC and log C_{max}) may provide useful information for BA and BE evaluation. There are different sources of variability for major PK parameters. Between-subject variability (ie, intersubject variability) is to describe the variability for the same product administered to different individuals. If the same product is given to the same individual multiple times (eg, in a replicate BE study), within-subject variability (intra-subject variability) can be calculated. A subject-by-formulation variability is used to describe the exposure change for the same individual upon product switching. Different types of variability provide information on product and patient characteristics; hence, the values should be assessed when needed.

In addition to major PK parameters, such as C_{max} and AUC, T_{max} is an important PK parameter obtained in a BA/BE study. T_{max} has relatively high between-subject variability and does not appear to follow common distributions, such as normal or log-normal distributions. No formal statistical testing has been recommended on the basis of T_{max}. However, T_{max} values can be important for drugs with quick treatment onset. Therefore, comparison of medians (and ranges) of T_{max} among different treatment groups are considered. There are cases where products are not deemed bioequivalent in which the median T_{max} are substantially different between the test product and the reference product, even if C_{max} and AUC satisfy the BE standard.

14.4.2 Statistical analysis for BE studies

14.4.2.1 Average BE

Average BE testing is the most commonly used, which is established to ensure that the true mean (ie, population average) difference of the major PK parameters between the test product and reference product is within a predefined range.[19] The test product is considered bioequivalent to the reference product only when the true means for all the major PK parameters (eg, AUC and C_{max}) of the test product is neither superior nor inferior to that of the reference product.

For most drugs with a reasonable therapeutic window, an approximately 20% change in mean major PK parameters (eg, AUC and C_{max}) is not anticipated to yield clinically meaningful changes in both efficacy and safety. Thus, for these products, the upper and lower thresholds (ie, δ_L and δ_U) are defined as ± 0.22 in log scale or ranging from 80% to 125% under normal scale. However, BE limits are not fixed for all products. For HV products (Section 14.5.2) and NTI drugs (Section 14.5.3), BE limits may be readjusted accordingly.

Average BE assessment is set up differently from a conventional statistical test aimed to demonstrate treatment difference between two groups. The null hypothesis for a conventional statistical test is the true means from the test product and the reference product are the same (ie, $\mu_T - \mu_R = 0$). The null hypothesis is rejected and alternative hypothesis is taken (ie, $\mu_T - \mu_R \neq 0$) when the calculated *P*-value based on the established statistic test is smaller than a predefined value (eg, 0.05). This predefined *P*-value is used to control type I error rate. For the sake of discussion, only in this paragraph, we call two products with equal true means of the PK parameters the "bioequivalent" products. Under the conventional statistic test, the type I error is to falsely claim "inequivalence" for two bioequivalent products. But this is not the focus of BE testing. Instead, BE testing should focus on falsely claiming BE for two inequivalent products. Therefore, different approaches are taken for average BE evaluation.

One approach is known as confidence interval approach. The confidence interval approach is to ensure the true mean difference of a major PK parameter between the test product and the reference product is likely to be sufficiently small. In this approach, the point estimate and 90% confidence interval of the mean difference of a specific PK parameter (eg, log AUC or log C_{max}) between the test product and the reference product can be calculated based on data. BE can be established if all estimated 90% confidence intervals of the major PK parameters are within the BE limits (eg, ± 0.22 in log scale or ranging from 80% to 125% under normal scale). As just discussed, the error that needs to be controlled in BE evaluation is falsely claiming BE for two inequivalent products, which is the Type II error for a conventional statistical hypothesis test. A Type II error cannot be directly controlled in a conventional hypothesis test. However, the Type II error can be reflected in confidence interval approach as the probability of not including an alternative value (ie, upper and lower BE limit). Therefore, BE can be concluded with an error rate not greater than 10% if the 90% confidence interval falls within the predefined BE limits.

The second approach is two one-sided tests (TOST). TOST ensures that the true mean difference of a PK parameter between a test product and a reference product is unlikely to be either too large or too small. TOST is performed with two sets of null and alternative hypotheses:

Test 1: Null hypothesis: $\mu_T - \mu_R \geq \delta_U$; alternative hypothesis: $\mu_T - \mu_R < \delta_U$

Test 2: Null hypothesis: $\mu_T - \mu_R \leq \delta_L$; alternative hypothesis: $\mu_T - \mu_R > \delta_L$

where μ_T and μ_R are the true means for the test and reference products, respectively; and δ_L and δ_U are the predefined regulatory upper and lower limits. Each test can be performed independently. If both null hypotheses are rejected, alternative hypotheses can be accepted and so can its intersection (ie, $\delta_L < \mu_{T-}\mu_R < \delta_U$). The null hypothesis for each of the tests is rejected from one side at a *P*-value of 0.05 level ($\alpha = 0.05$). Thus, the TOST approach is equivalent to the approach using 90% confidence interval.

14.4.2.2 Population BE and individual BE

In addition to the average BE test, two other statistical tests are recommended by the FDA's BE guidance—population BE test and individual BE test.[19] The focus for average BE assessment is to compare true mean difference between different treatment groups, whereas population BE and individual BE consider both the true mean difference and the variance difference.

Population BE is constructed to ensure prescribability. A physician may choose to prescribe to a new patient either a brand name product (innovator's product) or a generic product shown to satisfy population BE criteria. The general efficacy and safety profiles for the two products should be anticipated to be similar. To ensure prescribability, an equation is constructed to compare the expected squared difference between a test product (T) and a reference product (R) administered to two different patients (the *j*th and *j'*th patient) with the same reference product (R) given to these two

patients (Eq. (14.5)). In Eq. (14.5), μ_T and μ_R are the true means of the test product and reference product, respectively. σ^2_{TT} is the total variance for the test product, and σ^2_{TR} is the total variance for the reference product. The final test formula is shown in Eq. (14.6). It is important to ensure the calculated value is smaller than a predefined threshold in order to conclude BE of the test product.

$$\frac{E(T_j - R_{j'})^2}{E(R_j - R_{j'})^2} = \frac{(\mu_T - \mu_R)^2 + (\sigma^2_{TT} + \sigma^2_{TR})}{2 \cdot \sigma^2_{TR}} \tag{14.5}$$

$$\Theta_p = \frac{(\mu_T - \mu_R)^2 + (\sigma^2_{TT} + \sigma^2_{TR})}{\sigma^2_{TR}} \tag{14.6}$$

The point of the individual BE is to ensure switchability, which allows switching from a brand name product to a generic product without compromising the efficacy and safety in the patient. The comparison is between the expected squared difference of the mean PK variable for the same subject (the *j*th subject) receiving the test product (T) and the reference product (R), and that from the same individual receiving reference products (R) multiple times (Eq. (14.7)). As shown in Eq. (14.7), μ_T and μ_R are the true means of the test product and reference product, respectively. In addition, σ^2_{WT} and σ^2_{WR} are the within-subject variability (ie, intrasubject variability) for patients receiving the test product and reference product alone, respectively, and σ^2_D is a term known as subject-by-formulation variance, which describes the PK variable change for the same individual upon product switching. The final test formula for individual BE is shown in Eq. (14.8). The test statistics must be smaller than a predefined limit in order to claim individual BE. To allow full assessment of individual BE, a replicate BE study is necessary.

$$E = \frac{E(T_j - R_j)^2}{E(R_j - R'_j)^2} = \frac{(\mu_T - \mu_R)^2 + \sigma^2_D(\sigma^2_{WT} + \sigma^2_{WR})}{2 \cdot \sigma^2_{WR}} \tag{14.7}$$

$$\Theta_I = \frac{(\mu_T - \mu_R)^2 + \sigma^2_D(\sigma^2_{WT} + \sigma^2_{WR})}{\sigma^2_{WR}} \tag{14.8}$$

14.4.3 Data analysis for BA studies

PK data obtained from BA studies are generally summarized by descriptive statistics. In general, no formal statistical test is required for data obtained from BA studies. However, the mean ratio and the 90% confidence interval of a major PK parameter should still be reported.

14.5 SPECIAL TOPICS FOR BA AND BE ASSESSMENT

BA/BE assessment has been widely applied in various product development programs. In recent years, certain areas have been expanded with the evolving science. In the following sections, BE studies requiring pAUCs, BE studies for HV drugs, and BE studies for NTI drugs are discussed in detail.

14.5.1 BE studies requiring pAUCs

For some products, pAUC has been used as an additional PK variable in BA or BE assessment.[1] It is known as the area under the concentration time curve over a time interval of interest.[28] Eq. (14.9) defines the pAUC between t_1 and t_2 (ie, $pAUC_{t1-t2}$). For example, if there is a need to quantify drug exposure between 8 and 12 hours postdose, $pAUC_{8-12}$ should be obtained. In reality, a pAUC can be calculated based on trapezoidal rule. The area between two adjacent observed time points can be calculated from the area of trapezoid. Then the pAUC can be obtained by summing up all the segments within a time interval of interest. It can be shown that a pAUC represents the average concentration over the time interval of interest. Multiple pAUCs contain the information of the shape of the PK profile. For instance, the PK profile following the administration of a single dose of Ambien CR, a Zolpidem extended release tablet indicated for the treatment of insomnia, under fasting conditions, can be divided into two portions. In this case, $pAUC0_{1.5}$ and $pAUC1.5_t$ can be calculated separately: $pAUC0_{1.5}$ represents the early exposure and $pAUC1.5_t$ represents the late exposure. $AUC0_\infty$ represents the total zolpidem exposure. Overall, the two pAUCs define the shape of the underlying PK profile:

$$pAUC_{t1-t2} = \int_{t1}^{t2} C_t \cdot dt \tag{14.9}$$

The initial concept of pAUC was introduced back in the early 1990s. By then, scientists actively searched for an alternative PK variable to describe the rate of absorption. As we know, C_{max} is related to drug absorption rate and is used in BA and BE assessment. Unfortunately, C_{max} can be driven by other factors, such as the extent of drug absorption, distribution, and elimination. Chen et al. recommended using pAUC in BA and BE evaluation.[31] Tozer et al. interpreted pAUC as a measurement of early exposure.[32] In recent years, pAUCs have been recommended by the FDA to assess BA and BE for several products.

Also, pAUCs may serve several different roles in products with different release mechanisms. Some products follow the same release mechanism (monomodal release mechanism). For example, an immediate release tablet always releases the drug following zero-order or first-order kinetics. There is no release pattern change over time. Through simulations, it can be shown that pAUC, defined as the AUC between the time of dosing to the time of reaching the maximal concentration, is more sensitive than C_{max} in detecting difference in drug release for an immediate-release product. This pAUC may be correlated to clinical responses for some immediate-release products.[33] In some cases, conclusions of BE-based C_{max} and total AUC between the test product, and the reference product may be insufficient to demonstrate that there is no difference in safety or efficacy if the systemic concentration time profiles of the test product and the reference product are different (eg, time to reach peak drug concentration (T_{max}) is different). For example, differences in the shape of the systemic concentration profile between the test and reference products could imply that the test product may not produce the same clinical response as the reference product. The FDA's guidance recommends that pAUC be included as a measurement for BA/BE assessment.[1] In 2009, a workshop was sponsored by the American Association of Pharmaceutical Scientists (AAPS) in Baltimore, MD.[34] It was concluded by the experts from academia, industry, and regulatory bodies that "the current regulatory approaches criteria for BE evaluation were considered adequate for the assessment of therapeutic equivalence and interchangeability of conventional monophasic extended-release formulations." The current regulatory approaches refer to general C_{max} and AUC.

It has been found that pAUC is extremely useful for products containing both an immediate-release component and an extended-release component (multimodal release mechanism). In recent years, various products with multimodal release mechanisms were developed to ensure unique clinical needs, which are usually characterized as rapid onset, maintained pharmacological duration, and rapid offset. Multiple release characteristics can be altered to generate various PK profiles. For example, the drug release rate (Ka), and extent (F) for both immediate-release and extended-release components may be altered. In addition, the proportion of the immediate component and the delay time (T_{lag}) for the drug released from the extended-release component can be changed through formulation techniques. Hence, it is possible to generate different shapes of PK profiles with identical C_{max} and AUC, and pAUCs are critical to characterize BA and ensure BE.

A good example of this comes from the sleeping aid products mentioned previously. The Ambien CR tablet label indicates that Ambien CR (zolpidem tartrate) is an extended-release formulation of zolpidem with a bimodal release profile. Each Ambien CR tablet is comprised of an immediate-release component and an extended-release component, providing a rapid increase of zolpidem plasma concentration followed by a sustained plasma concentration level of zolpidem over the typical time interval for sleep.[35]

Thus, Ambien CR is a multiphasic modified-release formulation designed to meet the specific clinical needs for treating insomnia. According to the FDA-approved package insert, clinical studies showed statistically significant improvement of sleep induction in patients receiving Ambien CR over placebo. In addition, Ambien CR decreased wake time after sleep onset for the first 7 hours during the first 2 nights and for the first 5 hours after 2 weeks of treatment. As this multiphasic modified-release dosage form is designed to achieve both rapid onset of activity and sustained activity with sufficient duration, FDA suggests that additional BE metrics may be appropriate to ensure that a generic (test) version is therapeutically equivalent to the corresponding reference product. Thus, for Ambien CR, the pAUC metrics are proposed in addition to the traditional ($AUC_{0-\infty}$ and C_{max}) metrics.[36]

For certain locally acting drugs, pAUCs can also be important in establishing equivalence at the site for drug action. Take mesalamine, for example. The drug has been approved for the treatment of ulcerative colitis,[37] an inflammatory disease localized in the colon and presenting with open sores or ulcers. Pharmacologically, mesalamine is thought to act locally in the colon in order to control the inflammation, even though the drug by itself can be systemically absorbed. An effective treatment is on the basis of adequate delivery of mesalamine into the inflammation site in the colon. Based on the mechanism of action, a product with a delayed-release feature should be more efficacious than a product without delayed release, even though both products may yield the same C_{max} and AUC. In this case, pAUC is recommended to be used in the BE study to ensure that the delayed-release feature of an effective product is carried to its generic version, and further, to distinguish the mesalamine release exclusively in the colon versus mesalamine released partially outside the colon.[38]

14.5.2 BE evaluation for HV drugs

Sample size can be an issue for a BE study for a drug that is HV. HV drugs are drugs that exhibit at least 30% of the within-subject variability (intrasubject variability)

in one of the PK variables (eg, AUC or C_{max}).[39–41] There are data suggesting that more than 20% of generic products requesting for marketing approval in the United States between 2003 and 2005 are considered HV drugs.[40] The main concern for a HV drug is that a BE study requires a large sample size. For example, the general BE limits are 80% and 125%. Let us assume the underlying (ie, true) geometric mean ratio between the test product and reference product is 1. If the within-subject variability is 15%, only 10 subjects are needed in a two-way crossover study in order to ensure 80% of power to determine BE. The required sample size is tripled when the within-subject variability is 30%. Moreover, the sample size should be as large as 108 if the within-subject variability is 60%. If the underlying geometric mean ratio is not equal to 1, the required sample size should be further increased. For instance, if the underlying geometric mean ratio between the test product and reference product is 1.1 and the within-subject variability is 60%, the required sample size is 236 in a two-way crossover study with 80% power to demonstrate BE.[41]

Questions have been raised on the necessity of including large sample size in a BE study for a HV drug when the BE limits are set as 80% and 125%. Even though the cause of the high within-subject variability (ie, interoccasion variability) is unclear, some researchers believe that this variability is due to the drug rather than the formulation. Large presystemic metabolism, low BA, and acid liability appear to contribute to the large variability.[40,42] In addition, this large variability is also likely seen in the reference product from which the efficacy and safety information is derived. The large variability, if deemed to be clinically critical, is likely to yield efficacy and safety signals unfavorable to support regulatory approval of the reference product. In other words, the approval of the reference product implies clinical acceptance of the large within-subject PK variability and the product should have a wide therapeutic range.[43] Furthermore, the reference product may fail the BE testing with itself if the sample size is not large enough. Hence, it does not appear to be rational to enroll an unreasonably large sample size in a study with the objective to meet a fixed BE limit ranging from 80% to 125%.

The FDA has taken various actions to identify alternative BE assessment approaches for HV products. An advisory committee meeting was held in 2004. At the meeting, the committee recommended that FDA may expand the BE limits. One concept taken from individual BE testing was to allow reference scaling. The committee requested FDA to explore the feasibility of applying this approach to BE studies with drugs that are HV. In response to this request, the agency conducted several simulations, which show that reference scaling is a feasible approach for BE testing and the type I error can be well controlled.[44] The FDA's findings were presented to a second advisory committee meeting held in 2006.[45] Since 2006, FDA has formally adopted the reference-scaled average BE approach for drugs that are HV.[46]

The limits for a regular average BE testing are constant. In a regular average BE testing, the difference of the log-transformed true PK parameters ($\mu_T - \mu_R$) are supposed to be within a defined range (δ_L and δ_U) for a BE claim. Eq. (14.10) demonstrates this relationship. This relationship can be transformed into Eq. (14.11) because the upper and lower limits are symmetric around zero. The test product is considered bioequivalent to a reference product only when the 90% confidence interval of each PK variable of interest (eg, AUC and C_{max}) falls into the BE limit between 80% and 125%.[41]

$$\ln(0.8) = \delta_L \le (\mu_T - \mu_R) \le \delta_U = \ln(1.25) \tag{14.10}$$

$$(\mu_T - \mu_R)^2 \le \delta_U^2 = \ln(1.25)^2 \tag{14.11}$$

The key feature for a reference-scaled average BE assessment is to allow readjustment of the BE limits based on the within-subject variability of the reference product. Eq. (14.12) demonstrates the concept of a reference-scaled BE testing for a HV drug, in which σ_{WR}^2 is the variance term for within-subject variability of the reference product, and θ_s is the BE limit. The new BE limit, θ_s, is defined in Eq. (14.13), where the value of σ_{W0}^2 is a constant defined by the regulatory agency. Eq. (14.12) applies only when the within-subject variability of the reference product is greater than the predefined threshold value of variance, σ_{W0}^2. Eq. (14.14) is formed by combining Eqs. (14.12) and (14.13). For products whose within-subject variability of the reference product (σ_{WR}^2) is larger than the defined regulatory threshold for variance (σ_{W0}^2), the ratio of $\sigma_{WR}^2/\sigma_{W0}^2$ is always greater than 1. It means that the allowable limit is greater than ln(1.25). The larger the value of σ_{WR}^2, the greater the ratio ($\sigma_{WR}^2/\sigma_{W0}^2$). Hence, the allowed limits can be further expanded:

$$\frac{(\mu_T - \mu_R)^2}{\sigma_{WR}^2} \le \theta_s \tag{14.12}$$

$$\theta_s = \frac{[\ln(1.25)]^2}{\sigma_{W0}^2} = \frac{U_L^2}{\sigma_{W0}^2} \tag{14.13}$$

$$(\mu_T - \mu_R)^2 \le \frac{[\ln(1.25)]^2 \cdot \sigma_{WR}^2}{\sigma_{W0}^2} \tag{14.14}$$

For drugs whose within-subject variability (σ_{WR}) is less than or equal to σ_{W0}^2, BE testing is established on the basis of Eq. (14.15). In other words, BE testing uses

a constant average BE limit (ie, $\theta_s < \ln(1.25)^2/\sigma^2_{W0}$ for the mean difference of log-transformed PK parameters). Because σ^2_{W0} is a constant, the final BE testing can be shown in Eq. (14.16), which is the general form for average BE testing:

$$\frac{(\mu_T-\mu_R)^2}{\sigma^2_{W0}} \le \theta_s \frac{U^2_L}{\sigma^2_{W0}} = \frac{[\ln(1.25)]^2}{\sigma^2_{W0}} \tag{14.15}$$

$$(\mu_T-\mu_R)^2 \le U^2_L = [\ln(1.25)]^2 \tag{14.16}$$

Depending on the within-subject variability of the reference product, a different variance term (σ^2_{W0} or σ^2_{WR}) may be used in the test formula. As a result, the final BE limits may be expanded only when the within-subject variability of the reference product is large. Hence, this approach is also known as the *mixed-scaling approach*.[41]

To apply the mixed-scaling approach for a BE testing, within-subject variability of the reference product should be assessed. Thus, the study must be a replicate BE study (ie, a full or partial replicate design) where the reference product is administered to the same individual twice.[47,48] One key element in using mixed-scaling approach is to determine the cutoff value of the within-subject variability of the reference product, above which a reference-scaling is required. As defined by the FDA guidance, the practical cutoff value of the estimate of within-subject variability (σ_{WR}) is 0.294, even though the theoretical σ_{W0} is set as 0.25. Only when the estimate of within-subject variability for reference product is larger than 0.294 is a reference-scaling approach allowed. That is, the 90% confidence interval of the ratio for the PK parameter of interest should be within an increased BE range derived on the basis of the within-subject variability of the reference product. In addition, FDA guidance recommends that the point estimate of geometric mean ratio for a PK parameter should be within the regular BE limits of 80−125% as a secondary constraint.[46]

The mixed-scaling approach has been successfully applied to support approvals of multiple products. A good example is the postapproval formulation change for a mesalamine delayed-release tablet. Mesalamine has a high within-subject variability (>30%). The major PK parameters were calculated through noncompartmental analysis with data obtained from a single-dose full-replicate BE study. A reference-scaling approach was used and product passed the expanded BE limits, which led to the final approval of the product.[49]

14.5.3 BE evaluation for NTI drugs

In NTI drugs, relatively small changes in doses or exposures may yield serious treatment failure or severe adverse events. According to the Code of Federal Regulations (CFR 320.33), a narrow therapeutic ratio drug may be determined based on the following criteria:

1. There is less than a twofold difference in median lethal dose (LD_{50}) and median effective dose (ED_{50}) values.
2. There is less than a twofold difference in the minimum toxic concentration (MTC) and minimum effective concentration (MEC) in the blood.
3. Safety and effective use of the drug products requires careful titration and patient monitoring.[50]

In practice, some of the pharmacodynamics parameters, such as LD_{50}, ED_{50}, MTC, and MEC, are essentially unavailable from clinical trials or clinical practice.

The definition of an NTI drug was further discussed at the 2011 Advisory Committee (AC) meeting for Pharmaceutical Science and Clinical Pharmacology. Based on the AC recommendations and the description in the CFR, NTI drugs was defined as where small differences in dose or blood concentration may lead to serious therapeutic failures and/or adverse drug reactions that are life-threatening or result in persistent or significant disability or incapacity, such as warfarin.[51] Furthermore, it was considered that NTI drugs generally have the following characteristics:

- There is little separation between therapeutic and toxic doses (or the associated blood/plasma concentrations).
- Subtherapeutic concentrations may lead to serious therapeutic failure.
- They are subject to therapeutic monitoring based on PK or pharmacodynamic measures.
- They possess low-to-moderate (ie, no more than 30%) within-subject variability
- In clinical practice, doses are often adjusted in very small increments (less than 20%).

Two major approaches have been followed to establish BE for NTI drugs worldwide. The first approach is to tighten the BE limits[52] and the second approach is scaled-BE based on within-subject variability of the reference product. FDA uses the reference-scaled average BE approach for NTI drugs.[53,54]

FDA is recommending a full-replicate, crossover BE study for an NTI drug. In this study, the same subject receives both the reference product and the test product twice. This study design allows simultaneous comparisons of mean PK and within-subject variability between the reference product and the test product.

The comparison of means between a reference product and a test product is based on a reference-scaling approach similar to, but different from, what has been discussed in BE evaluation for HV drugs

(Section 14.5.2). A test product is considered bioequivalent to a reference product when the relationship shown in Eq. (14.17) is achieved. As discussed previously, μ_T and μ_R are the means of the log-transformed PK variables for the test product and the reference product, respectively. σ^2_{WR} is the within-subject variability of the reference product. θ^2_L is the upper BE limit for an NTI drug, which is narrowed down to ln(1.11). σ^{2*}_{WR} is the regulatory threshold. For an NTI drug,

$$\frac{(\mu_T-\mu_R)^2}{\sigma^2_{WR}} \le \frac{\theta^2_L}{\theta^{2*}_{W0}} = \frac{[\ln(1.11)]^2}{\theta^2_{W0}} \tag{14.17}$$

$$(\mu_T-\mu_R)^2 \le \frac{[\ln(1.11)]^2}{\theta^{2*}_{W0}} \cdot \sigma^2_{WR} \tag{14.18}$$

Eq. (14.18) is derived from Eq. (14.17) by multiplying σ^2_{WR} from both sides. In Eq. (14.18), the value of σ^{2*}_{W0} is 0.10 as defined by the regulatory agency. If a drug's within-subject variability (σ^2_{WR}) is also 0.10, $\sigma^2_{WR}/\sigma^{2*}_{W0}$ may cancel out. So, $(\mu_T-\mu_R)^2 \le [\ln(1.11)]^2$, which means the 90% confidence interval of the ratio of the major pharmacokinetic parameter between the test product and the reference product must be between 90% and 111%. The smaller the within-subject variability of the reference product, the narrower the BE range is. Likewise, if the within-subject variability of the reference product (σ_{WR}) is larger than 0.10, the allowable BE range can be greater than 90−111%. Under this situation, FDA requires that all PK parameters for BE testing must be within the regular BE range of 80−125% by using an unscaled average BE assessment.[55,56]

In addition to compare means of the PK parameters, the within-subject variability between the test product and reference product are compared in a BE testing. The main objective is to ensure the test product will not yield meaningfully larger within-subject variability than the reference product. The null hypothesis and alternative hypothesis are shown here[55]:

$$\text{Null hypothesis: } \frac{\sigma_{WT}}{\sigma_{WR}} > \delta$$

$$\text{Alternative hypothesis: } \frac{\sigma_{WT}}{\sigma_{WR}} \le \delta \tag{14.19}$$

The test statistics are used to compare the within-subject variability between the test product and reference product. Here, δ is the regulatory threshold, which can be considered as the nonsuperiority margin. If the confidence interval approach is taken, the 90% confidence interval of $\frac{\sigma_{WT}}{\sigma_{WR}}$ can be obtained through an F distribution. The upper value of the 90% confidence interval can be calculated by using Eq. (14.20), where S_{WT} and S_{WR} are the estimate of σ_{WT} and σ_{WR}, respectively; and ν_1 and ν_2 are the degrees of freedom for the F-test, where ν_1 is for the test product and ν_2 is for the reference product. The α value is set as 0.1. As shown in the FDA guidance on warfarin, to support a BE claim, the upper 90% confidence interval should be less than 2.5[55,56]:

$$\text{Upper 90\% confidence interval: } \frac{S_{WT}/S_{WR}}{\sqrt{F_{1-\alpha/2}(\nu_1,\nu_1)}} \tag{14.20}$$

Several product-specific BE guidances have been developed to guide the BE study design and data analysis for NTI drugs. For example, in 2012, the FDA published a BE guidance for warfarin. The guidance indicated that a full-replicate BE study is recommended and a detailed data analysis procedure has been provided.[56] The revised BE guidance for tacrolimus became available in 2012. The guidance also recommends a full-replicate BE study under both fast and fed conditions. In addition, the guidance requests the applicant to consider a scaled BE approach for data analysis.[57] These guidance documents laid out a pathway for companies to develop a generic product of an NTI drug.

14.6 BIOWAIVER AND BCS

In vivo studies are essential for BA and BE evaluation. However, as per FDA guidance, in the following three situations, the requirement for an in vivo study can be waived (ie, biowaiver):

- BA and BE for certain products are self-evident.
- In vitro data are sufficient to support BA and BE claims.
- Biowaiver may be granted based on BCS.

The details are discussed in the following sections.

14.6.1 BA and BE are self-evident

Biowaivers may be granted when the product is in solution because BA and BE are self-evident. A solution for parental, otic, or ophthalmic use does not require additional in vivo BA or BE evaluation if the active and inactive ingredients has the same concentration as another approved product. In addition, no in vivo BA/BE assessment for nasal solution, solution for aerosolization, nebulization, and skin use, or oral solution is necessary if the active ingredient is presented as the same concentration in the same dosage form as another approved drug product.[58] However, it is also critical to ensure that the inactive ingredient is not expected to alter the local or systemic absorption of the active drug.

14.6.2 BA and BE claim based on in vitro data

Sometimes in vitro data are sufficient to support biowaivers.[58] Biowaivers can be given to different strengths of the same solid dosage form based on information of formulation and in vitro dissolution test, provided that the following three conditions are met: (1) BE has been established on one of the strengths of the product; (2) the formulation of the to-be-waived strength product is proportionally similar to that of the other strength of product used in the BE study; (3) there is comparable dissolution of these two strengths of the products.

In vitro dissolution testing allows the direct comparison of the dissolution profiles of the test and reference products. Generally, a model-independent approach is recommended by the guidance. For example, similarity factor (f_2) can be calculated by using Eq. (14.21):

$$f_2 = 50 \cdot \log\left[100 \cdot \frac{1}{\sqrt[2]{1 + (1/n) \cdot \sum_{t=1}^{n} (R_t - T_t)^2}}\right] \quad (14.21)$$

In this equation, n is the total number of sampling time points. R_t and T_t are the cumulative percentage dissolved at the each of the selected time points from the reference product and the test product. The f_2 value between 50 and 100 is suggestive of similar dissolution profiles.[59]

In addition, biowaiver can be granted if in vivo and in vitro correlation has been established.

14.6.3 Biowaivers and BCS

Sometimes biowaivers may be granted by relying on BCS,[60] a scientific framework to categorize a drug substance based on its solubility and permeability (Table 14.3).

A BCS I product should show high solubility, rapid dissolution, and high permeability. In general, in vivo BE studies may not be required for a BCS class I product with a wide therapeutic index.[61] *High solubility* means that the drug in the highest strength is able to dissolve in no more than 250 mL of aqueous solution over the pH range of 1−6.8 at 37 °C ± 1 °C. The ionization characteristics of the test drug can be used to determine the number of pH conditions that should be studied. An immediate-release product with rapid dissolution is anticipated to release 85% of the labeled drug within 30 minutes, using US Pharmacopeia (USP) Apparatus I at 100 rpm or Apparatus II at 50 rpm in a volume of 500 mL or less in each of the following media: (1) 0.1 N HCl or simulated gastric fluid USP without enzymes; (2) a pH 4.5 buffer; and (3) a pH 6.8 buffer or simulated intestinal fluid USP without enzymes. An IR product is considered very rapidly dissolving when 85% or more of the labeled amount of the drug substance dissolves within 15 minutes using the abovementioned conditions. A highly permeable drug is determined as at least 85% of the administered dose can be absorbed, given that there is no instability issue of the drug in the GI tract.[61]

TABLE 14.3 Biopharmaceutical Classification System

BCS class	Solubility	Permeability
I	High	High
II	Low	High
III	High	Low
IV	Low	High

Multiple approaches can be followed to determine the permeability of a drug. PK data from human studies can be used to determine if a compound is highly permeable. For example, if absolute BA is reliably estimated (with adequate sample size and reliable study design and conduct) as 85% or above, the data by itself is sufficient to support that a compound is highly permeable. Mass balance study results based on isotope-labeled drugs can be used to estimate the extent of drug absorption. However, a mass balance study usually has large variability. The results alone might not be reliable. Intestinal permeability studies can be used to determine the permeability of a drug. These studies include in vivo human intestinal perfusion studies, in vivo or in situ intestinal perfusion studies in animals, in vitro permeation studies using excised human or animal intestinal tissues, or in vitro permeation studies across a monolayer of cultured epithelial cells.[61]

It should be noted that existence of efflux transporters may complicate the interpretation of the in vitro permeation study, and in vivo or in situ animal model results. The in vitro and in vivo or in situ animal model studies are suitable for passively transported drugs. If the cell or animal models do not reflect the transporter level in human, the classification for permeability may be misleading. To avoid this situation, the FDA guidance recommends that limiting the use of nonhuman permeability test methods for drug substances that are transported by passive mechanisms only. For instance, lack of dose/exposure-dependent change in BA in human studies may be suggestive that passive transport is the main mechanism for drug absorption.[61]

Stability in the GI tract is another factor should be taken into consideration when the permeability data is taken from total radioactivity from a mass balance study or from the loss of drug from an in vivo or in situ animal permeation study. In vitro stability test is

recommended to be conducted at 37°C over the time period that represents the appropriate contact time in the relevant fluids in body. For instance, stability study can be conducted over 1 hour in gastric fluid or 3 hours in intestinal fluid. A greater than 5% loss of the drug amount is indicative of instability of the drug.[61]

Biowaivers may be granted for a BCS III product. EMA[62] and Health Canada[63] both allow biowaivers for a BCS III product. In the latest draft FDA guidance published in 2015, a BCS III product may be considered for biowaivers.[61] This draft guidance is currently available for public comments.

BCS-based biowaivers provide a convenient way to avoid unnecessary in vivo BA and BE evaluation. It has been widely used to support NDA and ANDA submissions for various oral dosage forms.

14.7 SUMMARY AND FUTURE PERSPECTIVES

BA/BE evaluations are routinely performed in product development programs. BA/BE studies provide critical information to support product approval.

BA/BE assessment is typically based on a well-designed phase I clinical trial with PK samples collected at various time points from healthy volunteers or patients. The parent compound, and its metabolites when necessary, can be assessed through a valid bioassay. PK parameters, such as C_{max} and AUC, can be obtained through concentration-time profiles obtained from each individual. For special products with complicated release mechanisms, pAUC can be calculated over the time intervals of interest. Log-transformed major PK parameters between a test product and a reference product are compared. Average BE testing is the most common approach. It is designed to ensure the true mean difference between the log-transformed PK parameters between the test product and the reference product falls within a predefined range (ie, ln(0.8) to ln(1.25)). Two products are considered bioequivalent if the 90% confidence intervals of the ratios for all PK parameters of interest between the test product and the reference product are within the defined BE limits. However, the BE limits are not always fixed. For a HV product, the BE limits can be expanded based on the within-subject variability of the reference product. The approach is known as a reference-scaling approach. For an NTI product, the BE limits are defined by within-subject variability of the reference product together with unscaled BE approach.

In vivo BA/BE studies might not be necessary when the product can meet certain criteria. For example, in vivo BE studies for a BCS class I product (ie, high solubility, high permeability, and rapid dissolution) with wide therapeutic index may not be required. In addition, biowaiver can be given to the solution products when the BA is self-evident, as well as to different strengths of the same solid dosage form based on appropriate information on formulation and in vitro dissolution test.

The concept of BA/BE is still evolving. For example, long-acting injections, which can be given once every 2 weeks,[64] monthly,[65] or even longer,[66] are developed to provide convenient ways to administer drugs and improve compliance. These products usually employ relatively complicated release mechanisms. As a matter of fact, the extended dosing intervals require prolonged durations for BE studies, which can raise potential dropout issues and further complicate the interpretations of the results. In recent years, some novel concepts have been discussed to assess BA/BE for long-acting injections. For instance, Gehring and Martinez indicated that pAUC can be extended in the assessment of BA/BE for long-acting injections.[67] All these activities reflect the continuous efforts from the scientific and regulatory community to improve BA/BE assessment and hence protect public health.

References

1. FDA. Draft guidance for industry: bioavailability and bioequivalence studies submitted in NDAs or INDs—general considerations. <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm389370.pdf>; 2014 [accessed 21.05.15].
2. FDA. Guidance for industry: food-effect bioavailability and fed bioequivalence studies. <http://www.fda.gov/downloads/RegulatoryInformation/Guidances/UCM126833.pdf>; 2002 [accessed 21.05.15].
3. Zhang L, Wu F, Lee SC, Zhao H, Zhang L. pH-dependent drug-drug interactions for weak base drugs: potential implications for new drug development. *Clin Pharmacol Ther* 2014;**96**(2):266–77.
4. FDA. Guidance for industry: bioequivalence studies with pharmacokinetic endpoints for drugs submitted under an ANDA. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM377465.pdf>; 2013 [accessed 21.05.15].
5. Lindenbaum J, Mellow MH, Blackstone MO, Butler Jr VP. Variation in biologic availability of digoxin from four preparations. *N Engl J Med* 1971;**285**:1344–7.
6. Wagner JG, Christensen M, Sakmar E, Blair D, Yates JD, Willis 3rd PW, et al. Equivalence lack in digoxin plasma levels. *JAMA* 1973;**224**:199–204.
7. OTA. Drug bioequivalence. Recommendations from the Drug Bioequivalence Study Panel to the Office of Technology Assessment, Congress of the United States. *J Pharmacokinet Biopharm* 1974;**2**:433–66.
8. OTA. *Drug bioequivalence: a report of the Office of Technology Assessment, Drug Bioequivalence Study Panel*. Washington, DC: The Office: For Sale by the Superintendent of Documents, U.S. Government Printing Office; 1974.
9. Federal Register. Bioavailability and bioequivalence requirements; 1977.

10. FDA. Approved drug products with therapeutic equivalence evaluations. <http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/UCM071436.pdf>; 2013 [accessed 21.05.15].
11. Shargel L, Yu A. *Applied biopharmaceutics and pharmacokinetics*. 4th edn. Stamford, CT: Appleton & Lange; 1999.
12. 21CFR320.1 Bioavailability and bioequivalence requirements [Online]. <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?fr1/4320.1>; 2015 [accessed 21.05.15].
13. Fleisher D, Li C, Zhou Y, Pao L, Karim A. Drug, meal and formulation interactions influencing drug absorption after oral administration: clinical implications. *Clin Pharmacokinet* 1999;**36**:233–54.
14. Welling PG. Effects of food on drug absorption. *Annu Rev Nutr* 1996;**16**:383–415.
15. FDA. Guidance for industry: bioavailability and bioequivalence studies for nasal aerosols and nasal sprays for local action. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm070111.pdf>; 2003 [accessed 21.05.15].
16. FDA. Guidance for industry-topical dermatologic corticosteroids: in vivo bioequivalence. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm070234.pdf>; 1995 [accessed 21.05.15].
17. FDA. Clozapine U.S. package insert. <http://www.accessdata.fda.gov/drugsatfda_docs/label/2014/019758s073lbl.pdf>; 2014 [accessed 21.05.15].
18. FDA. Guidance on clozapine <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM249219.pdf>; 2005 [accessed 21.05.15].
19. FDA. Guidance for industry: statistical approaches to establish bioequivalence. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM070244.pdf>; 2001 [accessed 21.05.15].
20. El-Tahtawy AA, Jackson AJ, Ludden TM. Comparison of single and multiple dose pharmacokinetics using clinical bioequivalence data and Monte Carlo simulations. *Pharm Res* 1994;**11**:1330–6.
21. Singleton C. Recent advances in bioanalytical sample preparation for LC-MS analysis. *Bioanalysis* 2012;**4**(9):1123–40.
22. Thurman EM, Mills MS. *Solid phase extraction: principles and practice*. New York, NY: Wiley-Interscience Publication; 1998.
23. Gosetti F, Mazzucco E, Gennaro M, Marengo E. Ultra high performance liquid chromatography tandem mass spectrometry determination and profiling of prohibited steroids in human biological matrices. A review. *J Chromatogr B* 2013;**927**:22–36.
24. Jong LA, Uges DR, Franke JP, Bischoff R. Receptor-ligand binding assays: technologies and applications. *J Chromatogr B* 2005;**829**:1–25.
25. FDA. Guidance for industry: bioanalytical method validation. <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm368107.pdf>; 2013 [accessed 21.05.15].
26. FDA. Development of new stereoisomeric drugs. <http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm122883.htm>; 1992 [accessed 21.05.15].
27. FDA. Draft guidance on amphetamine aspartate; amphetamine sulfate; destroamphetamine saccharate; dextroamphetamine sulfate. <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm319975.pdf>; 2012 [accessed 21.05.15].
28. FDA. Meeting of the advisory committee for pharmaceutical science and clinical pharmacology. <http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/AdvisoryCommitteeforPharmaceuticalScienceandClinicalPharmacology/UCM207955.pdf>; 2010 [accessed 08.12.13].
29. DAILYMED: AMBIEN CR. <http://dailymed.nlm.nih.gov/dailymed/drugInfo.cfm?setid = 404c858c-89ac-4c9d-8a96-8702a28e6e76>; 2014 [accessed 18.08.15].
30. FDA. Draft guidance on Zolpidem extended release tablets. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM175029.pdf>; 2011 [accessed 18.08.15].
31. Chen ML. An alternative approach for assessment of rate of absorption in bioequivalence studies. *Pharm Res* 1992;**9**(11):1380–5.
32. Tozer TN, Bois FY, Hauck WW, Chen ML, Williams RL. Absorption rate vs. exposure: which is more useful for bioequivalence testing? *Pharm Res* 1996;**13**(3):453–6.
33. Dokoumetzidis A, Macheras P. On the use of partial AUC as an early exposure metric. *Eur J Pharm Sci* 2000;**10**(2):91–5.
34. Chen ML, Shah VP, Ganes D, Midha KK, Caro J, Nambliar P, et al. Challenges and opportunities in establishing scientific and regulatory standards for determining therapeutic equivalence of modified-release products: workshop summary report. *Clin Ther* 2010;**32**(10):1704–12.
35. FDA. Ambien CR ® U.S. package insert: <http://www.accessdata.fda.gov/drugsatfda_docs/label/2014/019908s035,021774s016lbl.pdf>; 2014 [accessed 21.05.15].
36. FDA. Guidance on zolpidem. <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm175029.pdf>; 2011 [accessed 21.05.15].
37. FDA. Asacol ® U.S. package insert: <http://www.accessdata.fda.gov/drugsatfda_docs/label/2015/019651s025lbl.pdf>; 2015 [accessed 21.05.15].
38. FDA. Draft guidance on mesalamine: <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm320002.pdf>; 2012 [accessed 21.05.15].
39. Blume H, Midha K. Bio-International 92, conference on bioavailability, bioequivalence, and pharmacokinetic studies. *J Pharm Sci* 1993;**82**:1186–9.
40. Davit BM, Conner DP, Fabian-Fritsch B, et al. Highly variable drugs: observations from bioequivalence data submitted to the FDA for new generic drug applications. *AAPS J* 2008;**10**:148–56.
41. Davit BM, Patel DT. *Bioequivalence of highly variable drugs. FDA bioequivalence standards*. New York, NY: AAPS Press and Springer; 2014. p. 139–64.
42. Cook J, Davit B, Polli J. Impact of biopharmaceutics classification system-based biowaivers. *Mol Pharm* 2010;**7**:1539–44.
43. Benet L. Why highly variable drugs are safer. FDA Advisory Committee for Pharmaceutical Sciences and Clinical Pharmacology meeting transcript. Available via US Food and Drug Administration Dockets. <http://www.fda.gov/ohrms/dockets/ac/04/transcripts/4034T2.htm>; 2004 [accessed 21.05.15].
44. Executive Secretary. Minutes of the meeting of the FDA Advisory Committee for Pharmaceutical Science and Clinical Pharmacology. Available via US Food and Drug Administration Dockets. <http://www.fda.gov/ohrms/dockets/ac/04/minutes/4034M1.htm>; 2004 [accessed 21.05.15].
45. Executive Secretary. Minutes of the meeting of the FDA Advisory Committee for Pharmaceutical Science and Clinical Pharmacology. Available via US Food and Drug Administration Dockets. <http://www.fda.gov/ohrms/dockets/ac/06/minutes/2006-4241m2.pdf>; 2006 [accessed 21.05.15].
46. US Department of Health and Human Services. Draft guidance for industry, bioequivalence of progesterone capsules. Available via Food and Drug Administration, Center for Drug Evaluation and Research. <http://www.fda.gov/downloads/Drugs/>

GuidanceComplianceRegulatoryInformation/Guidances/UCM209294.pdf>; 2012 [accessed 21.05.15].
47. Haidar SH, Davit BM, Chen M-L, et al. Bioequivalence approaches for highly variable drugs and drug products. *Pharm Res* 2008;**25**:237–41.
48. Haidar SH, Makhlouf F, Schuirmann DJ, et al. Evaluation of a scaling approach for the bioequivalence of highly variable drugs. *AAPS J* 2008;**10**:450–4.
49. Office of Clinical Pharmacology and Office of New Drug Quality Assurance Biopharmaceutics. Clinical pharmacology and biopharmaceutics review, NDA 204412. Available via Drugs@FDA. <http://www.accessdata.fda.gov/drugsatfda_docs/nda/2013/204412Orig1s000ClinPharmR.pdf>; 2012 [accessed 21.05.15].
50. 21CFR320.33. <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?fr=320.33>; 2015 [accessed 21.05.15].
51. Yu L, Jiang W, Zhang X, Lionberger R, Makhlouf F, Schuirmann DJ, et al. Novel bioequivalence approach for narrow therapeutic index drugs. *Clin Pharmacol Ther*. 2015;**97**(3):286–91.
52. European Medicines Agency. Guideline on the investigation of bioequivalence <http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2010/01/WC500070039.pdf>; 2010 [accessed 21.05.15].
53. US Food and Drug Administration. April 13, 2010 Meeting of the Pharmaceutical Science and Clinical Pharmacology Advisory Committee: topic 1, revising the BE approaches for critical dose drugs. <http://www.fda.gov/AdvisoryCommittees/Calendar/ucm203405.htm>; 2010 [accessed 21.05.2015].
54. Yu LX. Quality and bioequivalence standards for narrow therapeutic index drugs. <http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedand.Approved/ApprovalApplications/AbbreviatedNewDrugApplicationANDAGenerics/UCM292676.pdf>; 2011 [accessed 21.05.15].
55. Jiang W, Yu LX. *Bioequivalence for narrow therapeutic index drugs. FDA bioequivalence standards*. New York, NY: AAPS Press and Springer; 2014. p. 191–216.
56. FDA. Draft guidance on warfarin sodium. <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm201283.pdf>; 2012 [accessed 21.05.15].
57. FDA. Draft guidance on tacrolimus. <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm181006.pdf>; 2012 [accessed 21.05.15].
58. CFR 21 320.22. Bioavailability and bioequivanece requirements. <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfCFR/CFRSearch.cfm?fr1/4320.22>; 2015 [accessed 21.05.15].
59. FDA. Guidance for industry: SUPAC-MR: modified release solid oral dosage forms: scale-up and postapproval changes: chemistry, manufacturing, and controls; in vitro dissolution testing and in vivo bioequivalence documentation; 1997.
60. Amidon GL, Lennernas H, Shah VP, Crison JR. A theoretical basis for a biopharmaceutic drug classification: the correlation of in vitro drug product dissolution and in vivo bioavailability. *Pharm Res* 1995;**12**(3):413–20.
61. FDA. Draft guidance for industry: waiver of in vivo bioavailability and bioequivalence studies for immediate-release solid oral dosage forms based on a biopharmaceutics classification system. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM070246.pdf>; 2015 [accessed 21.05.15].
62. EMA. European Medicines Agency: guideline on the investigation of bioequivalence. <http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2010/01/WC500070039.pdf>; 2008 [accessed 21.05.15].
63. Health Canada. Biopharmaceutics classification system based biowaiver: draft guidance document. <http://www.hc-sc.gc.ca/dhp-mps/alt_formats/pdf/prodpharma/applic-demande/guide-ld/bcs_guide_ld_scb-eng.pdf>; 2012 [accessed 21.05.15].
64. FDA. Risperidal Consta ® U.S. package insert. <http://www.accessdata.fda.gov/drugsatfda_docs/label/2014/021346s052lbl.pdf>; 2014 [accessed 21.05.15].
65. FDA. Abilify Maintena ® U.S. package insert. <http://www.accessdata.fda.gov/drugsatfda_docs/label/2014/202971s003lbl.pdf>; 2014 [accessed 21.05.15].
66. FDA. Invega trinza ® U.S. package insert. <http://www.accessdata.fda.gov/drugsatfda_docs/label/2015/207946s000lbl.pdf>; 2015 [accessed 21.05.15].
67. Gehring R, Martinez M. Assessing product bioequivalence for extended-release formulations and drugs with long half-lives. *J Vet Pharmacol* 2012;**35**(Suppl 1):3–9.

This page intentionally left blank

CHAPTER

15

Predictive Biopharmaceutics and Pharmacokinetics: Modeling and Simulation*

Y. Yang, Y. Zhao, J.Z. Duan, P. Zhao, L. Zhao and X. Zhang

U.S. Food and Drug Administration, Silver Spring, MD, United States

15.1 INTRODUCTION

Biopharmaceutics is the study of the physical and chemical properties of a drug, and its dosage form, as related to the onset, duration, and intensity of drug action. It bridges in vitro characteristics with in vivo performance. The term *biopharmaceutical performance* refers to the influence of pharmaceutical formulation variables on in vivo performance.[1] In vivo performance can be assessed by pharmacokinetics (PK), pharmacodynamics (PD), or clinical end points, among which PK is the most frequently assessed in vivo performance. Predicting biopharmaceutical performance is a common interest of scientists working on drug product development and regulation because such predictions may reduce drug development cost and timelines. Scientists have been developing qualitative models (such as the biopharmaceutics classification system (BCS) and the biopharmaceutical drug disposition and classification system (BDDCS)), and quantitative models (such as a number of models that will be discussed in this chapter) to predict in vivo performance for various drug products.

Predicting oral absorption is difficult due to the complicated process involved. Briefly, when a solid dosage form is administered orally, it will disintegrate, dissolve, transport into enterocytes, be metabolized, and travel along the gastrointestinal (GI) tract. This process can be affected by formulation properties, drug substance properties, and GI physiology. The qualitative models, such as BCS and BDDCS, are classification systems based on a drug substance's solubility and permeability, and the extent of drug metabolism to predict drug absorption, disposition, and potential drug–drug interactions (DDIs) in the intestine and liver.[2] Details on BCS can be found in chapter "Oral Drug Absorption: Evaluation and Prediction" of this book.

The BDDCS is a modification of the BCS. In the BDDCS, a drug substance can be categorized in terms of solubility and the extent of metabolism: class I (high solubility, extensive metabolism), class II (low solubility, extensive metabolism), class III (high solubility, poor metabolism), and class IV (low solubility, poor metabolism). In addition, in the BDDCS concept, both uptake and efflux transporters in the intestine and liver are important in determining drug disposition by controlling absorption and bioavailability.[2] The qualitative models give a quick assessment of a drug's in vivo performance during development. In many cases, qualitative prediction is not adequate to address drug development and regulatory issues. Therefore, quantitative models are developed.

The objective of this chapter is to introduce quantitative models that can be used to predict biopharmaceutical PK performance for orally administered drug products. The outline of this chapter includes the following: (1) a general introduction to the conventional- and

*The opinions expressed in this report by the authors do not reflect the views or policies of the Food and Drug Administration (FDA).

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00015-7

mechanism-based PK models, (2) detailed discussions about quantitative predictions of absorption, distribution, metabolism, and excretion (ADME) using biopharmaceutics and PK modeling and simulation and its application in both drug product development and in regulatory assessment, and (3) challenges, opportunities, and future research directions.

15.2 MODELING AND SIMULATION APPROACHES FOR BIOPHARMACEUTICS AND PK

This section summarizes the most commonly adopted modeling and simulation methodologies for the prediction of the in vivo performance of solid oral dosage forms. These modeling strategies include the following: conventional compartment PK modeling, including population pharmacokinetic (popPK) modeling; and physiologically based pharmacokinetic (PBPK) modeling to predict drug ADME to support drug product development and various regulatory activities. Conventional PK and popPK models are used to estimate PK parameters and their variability using observed data sets. Sometimes it is difficult to interpret the results obtained from the conventional popPK model due to the lack of physiological framework. PBPK models take advantage of the physicochemical properties of drug substances, formulation properties of drug products, and physiological variability for certain human populations to predict PK performance of the drug product. The PBPK model can be developed with little or no human PK data, and, the predictions may be very valuable for further drug development.

15.2.1 Conventional compartment PK modeling and population PK modeling

Simplified conventional compartment PK models were developed to describe PK and ADME. Most widely used conventional compartment PK models are one-, two-, and three-compartment models with zero- or first-order absorption and linear or nonlinear elimination. The absorption is generally described by one or more absorption rate constants (k_a). The actual drug absorption process in a model may be described as first order, zero order, or a combination. The selection of either first-order or zero-order absorption is driven by observed data. Systemic drug absorption after oral administration of a solid dosage form is usually assumed to follow a first-order process. Distribution can be described by volume of distribution (V_d), and elimination can be described by clearance (CL).

The population PK approach is tightly connected with the conventional PK modeling in a way that both approaches are empirical and data driven, which is often referred as a *top-down* approach. The term *population PK* was coined by Lewis B. Sheiner and his longstanding statistician friend and coworker, Stuart Beal.[3] For the purpose of population PK modeling, nonlinear mixed-effects modeling is the most frequently referred approach.[4] Nonlinear mixed effects modeling during drug development is at the heart of a "Learn and Confirm" paradigm advocated by Sheiner.[5] The importance of variability is highlighted in population PK analysis than in traditional PK analysis. One main premise in population PK is that each individual is characterized by his or her own PK parameters; thus, inferences about the population of PK parameters is analogous to the population of patients and identifies covariates (eg, age, body weight, laboratory values, concomitant medications, other diseases, etc.) that correlate with PK variability.

Computing software tools of NONMEM (ICON), Pharsight Phoenix NLME (Pharsight), Kinetica (Adept Scientific), and Monolix are available for nonlinear mixed effects modeling.[4] Both the US Food and Drug Administration (FDA)[6] and the European Medicines Agency (EMA)[7] issued a specific, detailed Guidance on Population PK for Industry for the purpose of planning, analyzing, and reporting the data in a submission package. Decisions in drug development are increasingly being made on model-based population analyses.

15.2.2 Physiologically based PK modeling

Scientists intended to build a physiologically based model to describe PK many years ago, prior to the development of today's conventional PK models, but they did not make much progress due to the limitations of computational tools at that time. Physiologically based models to describe PK became feasible more recently with the development of computational tools. PBPK modeling is an approach in which the interactions of a drug with all components of the body are integrated, with the primary aims of permitting mechanistic insights into the global behavior of the system to be gained and of making meaningful extrapolations.[8] Prediction of human PK or ADME in early drug development using PBPK modeling may facilitate the selection and risk assessment of drug candidates before they are used in humans. PBPK models[9] also allow the full time-course of systemic drug and metabolite exposure to be evaluated realistically under all the

TABLE 15.1 Summarization of the Major Components in Physiologically Based Pharmacokinetic Modeling and Simulation

Structures	Human physiological or PK parameters	Drug substance and formulation parameters
Physiological compartment	Human body weight	Molecular weight (MW)
Anatomy	GI transit time	Logarithm of partition coefficient between *n*-octanol and water (log*P*)
Experimental design	Aqueous volume in each organ	Logarithm of the acid dissociation constant (pK_a)
Full model	Blood flow in each organ	Dose
Minimal model	Hepatic extraction ratio (ER)	Solubility
	Blood:plasma partition ratio	Precipitation time
	Protein binding	Particle density
	pH in GI lumen	Permeability
	GI fluid components	Particle size distribution

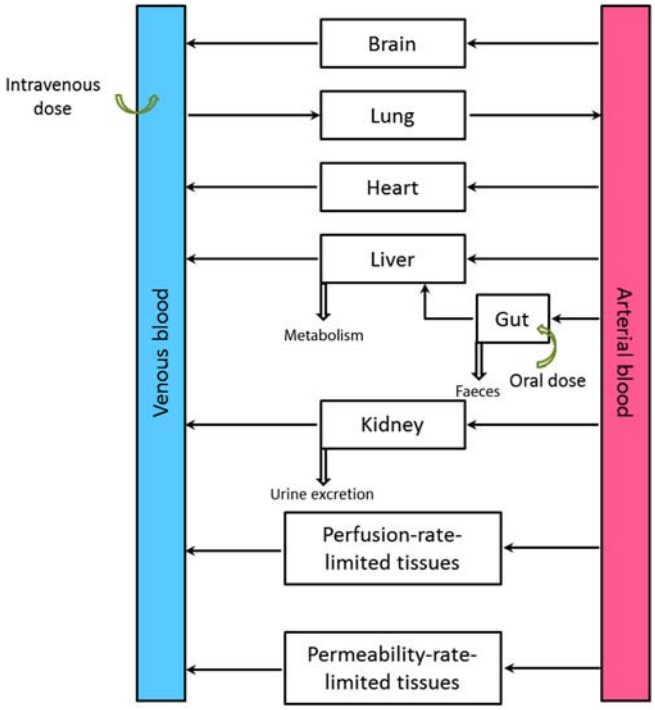


FIGURE 15.1 Schematic description of physiologically based pharmacokinetic modeling.

conditions and situations that are likely to be encountered during clinical development and beyond.

The concept of the PBPK model is to describe the concentration profile of a drug in various tissues, as well as in the blood over time, based on the drug characteristics, site and means of administration, and the physiological processes to which the drug is subjected. It comprises three major components: structural model, system-specific properties, and drug properties (listed in Table 15.1).[9] Structural model comprises the anatomical arrangement of the tissues and organs of the body, linked by perfusing blood. System-specific properties include organ mass or volume, blood flow, and tissue composition. Drug properties often include tissue affinity, plasma-protein-binding affinity, membrane permeability, enzymatic stability, and transporter activities. PBPK models are composed of many compartments corresponding to the different tissues of the body, connected by the circulating blood system (venous and arterial). A schematic of a PBPK model is shown in Fig. 15.1. A tissue can be described as either perfusion-rate-limited or permeability-rate-limited. Perfusion-rate-limited kinetics applies when the blood flow transporting the drug is often the limiting factor to distribution in the various cells of the body, while the tissue membranes present no barrier to diffusion.[10] A well-stirred tank has been the predominant flow-limited tissue compartment model in PBPK modeling.[11] Permeability rate-limited kinetics occurs where the permeability across the cell membrane becomes the limiting process.[10] When developing this kind of model, the mass-balance differential equations used in the model should describe both eliminating and noneliminating tissues.

The mass balance equation for the rates of change in drug concentration in a perfused noneliminating tissue is generally expressed as follows[12]:

$$V_t \frac{dC_t}{dt} = Q_t \times C_{art} - Q_t \times C_{vt} \tag{15.1}$$

where C_t is the total concentration in the tissue, V_t is tissue volume, Q_t is tissue blood flow, C_{art} is the drug concentration in the arterial blood reaching the tissue, and C_{vt} is the drug concentration in the venous blood leaving the tissue.

Accordingly, for a perfused eliminating tissue, the mass balance equation for the tissue can be expressed as follows[12]:

$$V_t \frac{dC_t}{dt} = Q_t \times C_{art} - Q_t \times C_{vt} - \text{in vivo } CL_{int,u} \times C_{vt,u} \tag{15.2}$$

where $CL_{int,u}$ is in vivo unbound intrinsic CL and $C_{vt,u}$ is the unbound venous tissue concentration.

Permeability rate-limited kinetics occurs where the permeability across the cell membrane becomes the limiting process.[10] Tissues possessing a permeability

barrier are modeled with permeability-limited equations that require consideration of two tissue subcompartments of vascular and extravascular spaces of noneliminating organs[13,14]:

$$\frac{dC_1}{dt} = \frac{Q_t}{V_1} * (C_{art} - C_1) - \frac{PS}{V_1} * \left(C_1 - \frac{C_2}{K_p}\right) \quad (15.3)$$

$$\frac{dC_2}{dt} = \frac{PS}{V_2} * \left(C_1 - \frac{C_2}{K_p}\right) \quad (15.4)$$

where C_1 and V_1 are the drug concentration and volume in the vascular subcompartment, C_2 and V_2 are the drug concentration and volume in the extravascular subcompartment, PS is the permeability surface area coefficient, and K_p is tissue:plasma distribution coefficient of the drug.

When selecting physiological compartments used in the model, specific considerations may be needed for a specific drug. Usually, liver can be considered as a separate compartment since drug metabolism occurs in the liver. For a lipophilic drug, fat can be considered as a separate compartment, given that the drug may have extensive partition into the fat. For a volatile drug, the blood compartment may be replaced by a blood−lung compartment.

During drug development, as more and more data become available, PBPK modeling can be used to predict the complexities of oral drug absorption, the likely intersubject PK variability. For small molecules, PBPK models have successfully been used to predict human exposure and to design the proposed dose regimens in phase I studies.[12] Intravenous and oral PK of 21 diverse compounds have been simulated.[15] With the use of Monte Carlo methods, a particular strength of PBPK modeling is the capability to predict variability in PK beyond observed limits and to anticipate individuals whose attributes combine to give extreme PK risk.[8,16,17]

Computing software tools such as GastroPlus (Simulations Plus Inc.), Simcyp Simulator (Simcyp Ltd), PK-Sim (Bayer Technology Services), MATLAB (MathWorks Inc.), and STELLA (isee systems) are available for PBPK modeling; note that the FDA does not endorse any computational software product or company. All these programs now strive to account for all relevant processes of the GI absorption of drugs, including release from the dosage form; decomposition/complexation in the GI tract; and the various mechanisms of drug uptake, efflux, and first-pass metabolism in either gut wall or liver; and to describe the interplay of these factors in determining the rate and extent of drug absorption from the GI tract.[18] The FDA has issued a guidance about using model-based strategy to evaluate and predict DDIs.[19] Recently, the EMA issued the specific detailed guideline on qualification and reporting of PBPK modeling and analyses.[20]

15.2.2.1 Absorption

Oral drug absorption is affected by drug substance properties (eg, dose, lipophilicity, pK_a, solubility, permeability, partition coefficient, and chemical stability), formulation factors (eg, particle size, release kinetics, and dissolution kinetics), and the physiology of the GI tract (eg, intestinal pH; transit time; GI motility; luminal and hepatic metabolism; blood flow; endogenous substances, such as bile salts; transporters; and exogenous substances such as nutrients). Permeability, solubility, and dissolution are the three primary drug- and formulation-specific factors that influence drug absorption.[21] Many details about GI physiology, drug permeability, solubility, and dissolution are described in chapter "Oral Drug Absorption: Evaluation and Prediction" of this book. Permeability reflects the physiological properties of membrane to the solutes. Solubility is one of the physiochemical properties of a drug substance that affects the absorption for a solid oral drug. Dissolution is the dosage form variable that determines the rate and extent of a drug dissolved in solution.

Passive permeability (P) of drug molecules across a cell membrane can be expressed as follows[22]:

$$P_{passive} = \frac{J}{\Delta C} = \frac{D \times K}{h} \quad (15.5)$$

where J is the flux (amount of material flowing through a unit cross section), ΔC is the drug concentration difference between the donor and receiver compartment, D is the diffusion coefficient, K is the partition coefficient, and h is the thickness of the cell membrane. The diffusion coefficient (D) depends on the drug molecular weight or size, and K is a measure of the solubility of the drug molecules in lipids. Therefore, passive permeability is related to membrane properties and drug properties. For a specific drug substance, passive membrane permeability should be a constant (P_m) that is independent of drug concentration.

The permeability for active transport can be presented similarly[22]:

$$P_{active} = \frac{J}{C} = \frac{J_{max} \times C}{K_m + C} \times \frac{1}{C} = \frac{J_{max}}{K_m + C} \quad (15.6)$$

where J_{max} is the maximum drug flux, K_m is the drug affinity to the carrier, and C is the concentration in the donor compartment. Obviously, active permeability depends on drug concentration.

Therefore, the total effective permeability depends on the concentration of drugs that are absorbed

through both passive diffusion and active transport, and it can be expressed as follows:

$$P_{\text{eff}} = P_{\text{passive}} + P_{\text{active}} = P_{\text{m}} + \frac{J_{\max}}{K_{\text{m}} + C} \tag{15.7}$$

When at very low drug concentrations (ie, $C \ll K_{\text{m}}$), drug permeability is independent of drug concentration. On the other hand, at high drug concentrations (ie, $C \gg K_{\text{m}}$), drug permeability depends on drug concentration.

Factors such as solubility, permeability, pK_{a}, and luminal pH behave interactively, affecting oral drug absorption. Ionization and pH plays an important role in the balance of solubility and permeability. The ionized form is usually more soluble in aqueous conditions than the neutral form, but the neutral form penetrates the enterocyte layer more easily by passive diffusion than the ionized form does.

Drug dissolution must take place with a solid drug dosage form in the GI tract before it can be absorbed. Factors that affect dissolution also have an impact on the fraction absorbed. The Noyes–Whitney equation can be used to describe the dissolution rate as follows[23]:

$$\frac{d_m}{d_t} = A \times \frac{D}{h} \times (C_{\text{s}} - C) \tag{15.8}$$

where d_m/d_t is the rate of the dissolution of a solid drug, A is the surface area of the interface between the dissolving substance and the solvent, D is the diffusion coefficient, h is the thickness of unstirred boundary layer, and C_{s} and C represent the solubility concentration of the drug substance and the concentration of the dissolved substance at a given time t, respectively.

As per the official definition (CFR title 21, section 320.1), bioavailability (F) means the rate and extent to which the active ingredient or active moiety is absorbed from a drug product and becomes available at the site of action. For orally administered drug products intended for systemic action, F is determined by a number of processes:

$$F = f_{\text{a}} \times f_{\text{g}} \times f_{\text{h}} \tag{15.9}$$

where f_{a} is the fraction of drug dose absorbed across the apical cell membrane into the cellular space of the enterocyte, f_{g} is the fraction of drug dose escaped from intestinal first-pass metabolism, and f_{h} is the fraction of the drug dose escaped from hepatic first-pass metabolism.

15.2.2.1.1 Compartmental absorption and transit model

PBPK modeling to predict oral drug absorption of solid drug dosage formulation has wide applications. Initially, a compartmental absorption and transit (CAT) model that divides the GI tract into seven compartments was developed.[24] The CAT model was developed by considering gastric emptying, small intestinal transit, in vivo dissolution, and first-order absorption.[24] It is a dynamic, mechanistic model that simulates and predicts the rate and extent of drug absorption from the human GI tract. The drug solution transfers from one compartment to the next in a first-order fashion. The colon is considered only as a reservoir. The CAT model is developed based on the following assumptions: (1) absorption from the stomach and colon is insignificant compared with that from the small intestine; (2) transport across the small intestinal membrane is passive; (3) dissolution is instantaneous; and (4) drug substance moving along the small intestine can be viewed as a process of flowing through a series of segments.[24] The basic equation for the CAT model is described as follows[25]:

$$\frac{dD_n}{dt} = k_{\text{t}} \times D_{n-1} - k_{\text{t}} \times D_n - k_{\text{a}} \times D_n \tag{15.10}$$

where D_n is the percent of dose in the nth GI compartment, n is the number of total compartments, k_{t} is the first-order transit rate constant, and k_{a} is the first-order absorption rate constant. Therefore, in the CAT model, the fraction of dose absorbed, f_{a}, can be estimated by[24]

$$f_{\text{a}} = 1 - \left(1 + \frac{k_{\text{a}}}{k_{\text{t}}}\right)^{-7} = 1 - (1 + 0.54 P_{\text{eff}})^{-7} \tag{15.11}$$

where P_{eff} is the human effective permeability, expressed in cm/hour.

The value of $P_{\text{eff, human}}$ can be obtained by the in situ intestinal perfusion method.[26] The model of this method was described in great detail in chapter "Oral Drug Absorption: Evaluation and Prediction." Briefly, the drug concentration is measured at the inlet and outlet of the perfusion tube. The drug permeability ($P_{\text{eff, human}}$) is calculated with the following equation[27]:

$$P_{\text{eff,human}} = \frac{Q \times \frac{C_{\text{in}} - C_{\text{out}}}{C_{\text{out}}}}{2\pi r L} \tag{15.12}$$

where Q is perfusion flow rate, C_{in} is inlet drug concentration of the perfusion tube, C_{out} is outlet drug concentration of the perfusion tube, and $2\pi rL$ is the mass transfer surface area within the intestinal segment, which is assumed to be the area of a cylinder with length L and radius r.

15.2.2.1.2 Advanced compartmental absorption and transit model

The advanced compartmental absorption and transit (ACAT) model was developed based on the

CAT model to include additional functions, including gut metabolism and modeling of modified-release drug products. This model is featured with first-order transport from one compartment to the next, and nonlinear metabolism/transport kinetics, nine GI compartments (stomach, duodenum, jejunum 1, jejunum 2, ileum 1, ileum 2, ileum 3, ileum 4, and colon), and six states of drug component (unreleased, undissolved, dissolved, degraded, metabolized, and absorbed), and three states of excreted material (unreleased, undissolved, and dissolved).[25] The ACAT model in predicting oral drug absorption considers drug physicochemical factors (pK_a, solubility, and permeability), dosage formulation factors (dose, particle size, and density), and physiological factors (gastric emptying, intestinal transit rate, first-pass metabolism, luminal transport, GI blood flow, and food effects). The GastroPlus software was developed based on the ACAT model. It is recommended that experimental data (human P_{eff}, solubility, etc.) be included to increase the reliability of GastroPlus. Since the determination of human P_{eff} is complex and expensive, in situ permeability values of rat or dog intestine are used instead. The human P_{eff} is estimated based on the measured permeability in rat or dog using a built-in correlation.[24] The data from epithelial cell culture models such as Caco-2 and MDCK can also be used for the purpose of prediction.[28]

15.2.2.1.3 Advanced dissolution, absorption, and metabolism model

Further, the advanced dissolution, absorption, and metabolism (ADAM) model was also developed based on the CAT model to include drug dissolution, GI fluid transit, drug degradation, gut wall permeation, intestinal metabolism, and active transport. The ADAM model consists of seven compartments (duodenum, jejunum I, jejunum II, ileum I, ileum II, ileum III, and ileum IV) for the small intestine. In each segment of the small intestine, ordinary differential equations are used to describe the dynamics of the amount of solid drug mass available for dissolution, the amount of solid mass trapped in the formulation and not available for dissolution, the amount of dissolved drug, and the drug concentration in the enterocyte.[29] The ADAM model considers drug factors (solubility and dissolution, drug stability, and permeability), dosage formulation factors (drug release and particle size), and physiological factors (gastric emptying, intestinal transit time, GI fluid pH and dynamics, first-pass metabolism, luminal transport, GI blood flow, and food effects).[29] The Simcyp software employs the ADAM model.

15.2.2.2 Distribution

The volume of distribution of a drug is a parameter that relates the amount of the drug in the body to the concentration of the drug at the site of measurement. For example, the volume of distribution at steady state (V_{ss}) in plasma corresponds to the equivalent plasma volume in which a drug is apparently distributed into the body at given doses. Physiologically, the volume of distribution at steady state is determined by[30]

$$V_{ss} = V_p + \sum t\left(V_t \times K_{t:p}\right) + V_e \times \left(e{:}p\right) \qquad (15.13)$$

where V_p, V_t, and V_e are the volume of plasma (about 4 L in humans), tissue, and erythrocyte, respectively; $K_{t:p}$ is the tissue-to-plasma partition coefficient; and $e{:}p$ is the erythrocyte-to-plasma ratio. The tissue-to-plasma partition coefficient ($K_{t:p}$)[12] can be calculated by the Poulin–Theil method using tissue composition-based equations[31] with the correction[32] or by the Rodgers–Rowland method using mechanistically derived equations.[33,34]

In addition, the PBPK modeling approach is used to predict local exposure to the brain of central nervous system–targeted drugs in humans from preclinical data. In one study,[35] acetaminophen was used as a test compound, as it is not subjected to active transport processes. The regional brain PK of acetaminophen was described by developing a physiologically based model.

Physiological volumes in humans involve plasma water (about 4 L), extracellular fluid (about 15 L), and total body water (about 40 L). Drugs bind to plasma and tissue protein. The binding of drugs to plasma proteins and tissue macromolecules exerts an important influence on the pattern of drug distribution. Plasma protein binding tends to restrict drug in the plasma, and tissue binding tends to accumulate drug in the tissue. Drug concentrations vary from tissue to tissue and is the highest where binding is most extensive. Overall, drug distribution is driven primarily by the passive diffusion of a drug along the concentration gradient created by the unbound drug in the plasma and tissues. Uptake and efflux transporters may also be involved and may either promote or limit a drug's distribution beyond that predicted by passive diffusion alone. Albumin, α1-acid glycoprotein, and lipoproteins are the plasma proteins principally involved in the binding of drugs. Albumin is the most abundant and has a concentration of about 40 g/L. Many drugs bind to albumin, particularly weak acids and neutral drugs. Drug binding to plasma or tissue proteins is a capacity-limited process that is governed by the law of mass action. Several factors decide drug protein

binding, including affinity of the drug for the protein, drug concentration if saturation is possible, and level of binding protein.

15.2.2.3 First-pass intestinal metabolism

In the human intestine, various metabolic enzymes (in particular, cytochrome P450s (CYPs) and uridine 5′-diphosphate-glucoronosyltranferase (UGT)) are expressed intensively[36] in epithelial cells, which could largely affect the bioavailability of orally administrated drugs. For example, CYP3A accounts for 80% of total intestinal P450s[37] and CYP3A substrates (eg, cyclosporine, nifedipine, midazolam, and verapamil) have their bioavailability reduced due to the first-pass metabolism in the gut wall.[36,38] Approximately one-third of the marketed drugs are metabolized to a substantial extent by CYP3A.[39] The impact of intestinal metabolism and intestinal bioavailability merits investigation for orally administrated drugs.

For the purpose of estimating the fraction of dose escaped intestinal first-pass metabolism (f_g), an approach named the "Q_{gut} model" has been developed.[40] This model treats the entire small intestine as a single homogenous organ following the form of the well-stirred liver model. It includes a flow term (Q_{gut}), which accounts for both permeability through the enterocyte membrane and the villous blood flow. Modeling equations of "Q_{gut} model" include[40]

$$f_g = \frac{Q_{gut}}{Q_{gut} + (fu_{gut} \times CL_{int,gut})} \tag{15.14}$$

$$Q_{gut} = \frac{CL_{perm} \times Q_{ent}}{CL_{perm} + Q_{ent}} \tag{15.15}$$

where fu_{gut} is the unbound fraction of drug in the enterocytes, $CL_{int,\ gut}$ is unbound intrinsic gut CL, CL_{perm} is permeability CL through the enterocytes, and Q_{ent} is mucosal blood flow (18 L/hour in an average human being). The value of $CL_{int,\ gut}$ is calculated by the product of the in vitro intrinsic CL and a correction factor taking into account enzyme content involved in substrate transformation. The value of CL_{perm} can be determined by the measured Caco-2 permeability at pH 7.4 and an estimate of the human small intestinal surface area.[41]

The predictive power of the Q_{gut} model has been further assessed in a study of 25 CYP3A substrates ranging over the full range of f_g values (0.07−0.94).[42] By using the Q_{gut} model, good agreement between predicted and in vivo f_g was noted for drugs with low to medium intestinal extraction when in vivo f_g is $\geq$0.5, but a lower predictive performance is observed when f_g is $\leq$0.5.[43]

15.2.2.4 Hepatic and renal CL

For solid oral drug formulation, after entering the systemic circulation, both hepatic metabolism in liver and renal excretion through the kidney can be involved in the drug's CL, hence determining the exposure levels of drugs and metabolites. A drug's total CL in the entire body equals the sum of the CLs of various organs[44]:

$$CL_{total} = CL_m + CL_r + CL_{others} \tag{15.16}$$

where CL_m is metabolic CL in liver, CL_r is renal CL, and CL_{others} is drug CL in other organs.

A drug's CL is the irreversible removal of it from certain organs, and is defined as[44]

$$CL = Q \times ER \tag{15.17}$$

where Q is organ blood flow and ER is the extraction ratio of the drug in the organ.

ER is an index of how efficiently the organ extracts drug from the blood that goes through it, and it ranges from 0 to 1. ER across an organ depends on the fraction of unbound drug in the blood (f_u) and intrinsic CL of the organ (CL_{int}, the inherent ability to remove the drug by the organ in the absence of any limitations) and is calculated by the following equation[44]:

$$ER = \frac{f_u \times CL_{int}}{(Q + f_u \times CL_{int})} \tag{15.18}$$

Drugs with low ER ($ER < 0.3$) show a capacity-limited CL ($CL = f_u \times CL_{int}$), where the CL is primarily determined by the unbound fraction of the drug and the intrinsic CL. Drugs with high ER ($ER > 0.7$) show organ blood flow-limited CL ($CL = Q$), where CL is directly affected by changes in blood flow.

15.2.2.4.1 Drug hepatic CL

The liver is a major site of drug elimination. Quantitatively, the liver normally accounts for the metabolism and the CL of a majority of drugs. Drug-metabolizing enzymes mainly localized to cellular endoplasmic reticulum catalyze the biotransformation of drugs to often pharmacologically inactive or sometimes active metabolites (eg, codeine (prodrug) conversion to morphine (active) and morphine conversion to morphine-6-glucuronide (active)) in humans. NADPH-cytochrome P450 reductase (P450R)/cytochrome P450s (P450s, namely CYP1A2, CYP2A6, CYP2C9, CYP2C19, CYP2D6, CYP2E1, and CYP3A4/5) are enzymes that mediate phase I oxidative reactions, whereas the uridine 5′-diphosphate-glucoronosyltranferase (UGTs, namely UGT1A1, UGT1A4, UGT2B4, and UGT2B7), sulfotransferases (SULTs, namely SULT1A1, SULT1A2, SULT2A, and SULT2B), glutathione *S*-transferases, and *N*-acetyltransferases are enzymes that mediate the

phase II conjugative reactions of most drugs.[44] For most drugs on the market, hepatic CL determines the drug exposure after the drug is absorbed.

Many mathematical models have been developed to describe in vivo hepatic drug CL. The most commonly used hepatic CL models are the venous equilibrium (well-stirred) model, the undistributed sinusoidal (parallel-tube) model, and the dispersion models.[45] The well-stirred model considers the homogenous drug concentration throughout the liver, and it assumes bulk flow, with the model equation as follows[45]:

$$CL_h = Q_h \times \frac{f_u \times CL_{int}}{(Q_h + f_u \times CL_{int})} \tag{15.19}$$

where Q_h is hepatic blood flow and CL_{int} is the intrinsic CL in the liver (see Eqs. (15.17) and (15.18)).

The parallel-tube model, which considers the liver to be a tube, assumes an exponential decline in drug concentration through the liver under the assumption of plug flow, with the following model equation[45]:

$$CL_h = Q_h \times (1 - e^{-(f_u \times CL_{int})/Q_h}) \tag{15.20}$$

The dispersion model is an intermediate of the previous two models.

For model building, it is necessary to investigate each determinant of the drug CL in vitro. The drug in vivo intrinsic CL (unit: L/minute or L/hour) can be extrapolated based on its in vitro intrinsic CLs based on the in vitro hepatocyte, liver microsome, or expressed enzyme experiments (unit: μL/minute per functional unit of the system).[46] In vitro experiments for CL prediction contain three elements: a general setup, a biological component, and a compound to be tested. The setup involves an incubation vessel and a suitable solution, but these nonbiological components may have a considerable effect on biological components and on the compound and its movement and fate through the system. The biological component involves the biological material, such as the enzyme sources (eg, liver slices, cultured hepatocytes, subcellular fractions (liver microsomes and homogenates), and recombinant expressed enzymes), and in vitro kinetics (eg, metabolic/elimination rate and competence, binding, and uptake/efflux factors (passive diffusion versus transporters)). For the compound, one should be able to measure its concentration in various compartments of the system and its biotransformation. In essence, meaningful information can be extracted from an in vitro system for in vitro–in vivo extrapolation.

Enzyme kinetics generally considers that 1 mol of drug interacts with 1 mol of enzyme to form an enzyme–drug intermediate. This rate process can usually be described by the Michaelis–Menton equation, which assumes that the rate of enzymatic reaction depends on the concentration of both enzyme and drug and that an enzyme–drug intermediate is initially formed, followed by the formation of the product and then the regeneration of the enzyme. The velocity of metabolism is described by the Michaelis–Menton equation:

$$v = \frac{V_{max} \times S}{S + K_m} \tag{15.21}$$

where v is the velocity of the reaction, V_{max} is the maximum velocity, S is the drug/substrate concentration, and K_m is the Michaelis constant, which is the substrate concentration when the velocity of the reaction is equal to 50% of the maximum velocity. The two parameters V_{max} and K_m determine the profile of a simple enzyme reaction rate at various drug concentrations. The relationship of the rate of metabolism to the drug concentration is nonlinear, which can be represented by a hyperbolic profile.

Scaling factors include the number of hepatocytes, liver microsomal content, expressed enzyme amount per gram of liver, liver weight, and body weight. A body of information on the prediction of CL by the metabolic enzymes other than CYP and UGT, particularly involving conjugation reactions mediated by amidases and sulfatases, is currently lacking.

15.2.2.4.2 Drug renal CL

The renal CL of a drug includes the combination of glomerular filtration, secretion, and reabsorption through proximal tubules. Using a noncompartmental approach, renal CL can be expressed as[44]

$$CL_r = \frac{A_u}{AUC_0^\infty} \tag{15.22}$$

where A_u is the drug amount ultimately excreted unchanged in urine and AUC_0^∞ is the area under the plasma concentration time curve.

Physiologically, renal CL of a drug equals

$$CL_r = \frac{\text{excretion rate}}{\text{plasma concentration}} = \frac{\text{filtration rate} + \text{secretion rate} - \text{reabsorption rate}}{\text{plasma concentration}} \tag{15.23}$$

When the drug is not actively secreted and the reabsorption is negligible, the renal CL of a drug is often related to the renal glomerular filtration rate. Drugs are usually filtered through the glomerulus in a nonselective process if they are in a free, nonprotein-bound form.

15.3 APPLICATION OF BIOPHARMACEUTICS AND PK MODELING AND SIMULATION IN DRUG DEVELOPMENT

In the previous section, we have discussed theoretical models that have been developed to predict the ADME of drug products. Numerous research articles have been published on PBPK modeling. A quick search using "physiologically based pharmacokinetics" as the keyword would find 1895 articles as of May 2015. Table 15.2 includes some representative case examples where mechanism-based models have been used in the

TABLE 15.2 Examples of Published Absorption PBPK Modeling Used in the Development of Solid Oral Dosage Formulation Within the Last 10 Years (2006–15)

Modeling purpose	Drug	Human BCS classification	Program used for modeling	References
Absolute oral BA	164 drugs	I–IV	CAT	47
PK prediction	Multiple compounds		Simcyp & ADAM, Gastroplus, PK-Sim, and MATLAB	48–52
Food effects	Novartis compounds	I, II, IV	GastroPlus & CAT	53–55
Interplay of food, drug, and CYP3A enzymes	Nifedipine	II	PK-Sim, Simcyp & ADAM	56,57
Oral absorption & food effects	Aprepitant	IV	STELLA	58
Food effects on drug absorption	Merck compound A	IV	STELLA	59
Food effects on absorption	Celecoxib	II	STELLA	60
Predicting PK for a buffered sublingual formulation	Propranolol	I	Gastroplus & ACAT	61
Impact of nonsink intestinal permeation on BA for poorly soluble compounds	1,3-dicyclohexyl urea	–	GastroPlus	62,63
Identify biorelevant solubility	Abbott compound	–	GastroPlus	64
Salts selection	Phenytoin	II	GastroPlus	65
Formulation optimization by PBPK/PD modeling	Nifedipine	II	Simcyp	66
Formulation optimization	SNX-2112	–	MATLAB	67
Formulation impact on oral absorption and CYP3A4 mediated gut metabolism	Multiple compounds	I–III	Simcyp	68
Formulation assessment	Atazanavir	II/IV	STELLA	69
Formulation development, PK profiles for immediate release and modified release products	Furosemide	IV	STELLA	70
In vivo relevant dissolution	Diclofenac		STELLA	71
Pharmaceutical product development	Multiple compounds	I–III	GI-Sim	72
PK of immediate release and extended release formulations	Cimetidine	III	PK-Sim	73
Formulation optimization (particle size)	Cilostazol	II	PK-Sim	74
In vivo PK	Commercial and Pfizer compounds	I–II	PK express model	75
Predict PK of various dosage forms and formulation	Melatonin	II	MATLAB Simulink	76
Predict PK of various formulations in various population	Quetiapine	II	Simcyp	77
In vivo relevant dissolution	Diclofenac	II	STELLA	78
Human intestinal first-pass metabolism	Multiple compounds	I–II	MATLAB	41
Intestinal metabolism	Verapamil/ compound A	–	MATLAB	79

area of oral absorption or predicting biopharmaceutical performance for oral dosage forms within the last 10 years.

Scientists have been very creative in utilizing absorption PBPK models to address scientific questions during drug development. As shown in Table 15.2, the drug substances that have been modeled covered a wide range of physicochemical properties, including BCS I–IV drugs. Computational platforms that have been used for modeling included Gastroplus, GI-Sim, MATLAB, PK-Sim, Simcyp, and STELLA. A direct use of these model types is to predict oral bioavailability and PK profiles for various drug substance and formulations.[47–52,73,76,77] As a practical tool for drug product development, oral PBPK absorption models have been used extensively in pharmaceutical development.[72] For example, it was reported to be used in salt selection,[65] particle size optimization,[74] optimization for PK profiles,[70] and even PD profiles.[56] Absorption PBPK models have also been used to study the intestinal metabolism.[41,56,57,66,68,79] Recently, efforts have been made in the area of quantitative prediction of food drug interaction using oral absorption PBPK models.[53–56,58,60] Scientists have been working to develop in vivo relevant dissolution methodology for a long time. With the application of mechanistic absorption models, researchers were able to link in vitro dissolution with in vivo performance in a mechanistic framework.[59,60,64,78] More important, mechanistic oral absorption models have been used to test scientific hypotheses and address barriers encountered in drug development. For example, nonsink permeation in the small intestine was incorporated into an absorption PBPK model to explain the low bioavailability, which could not be explained by low solubility alone.[62,63] The simulation was verified in rat studies by perfusing the drug substance to reach high plasma concentrations to mimic the nonsink permeation conditions.[63]

15.4 APPLICATION OF BIOPHARMACEUTICS AND PK MODELING AND SIMULATION IN REGULATORY ACTIVITIES

15.4.1 In the new drug evaluation

There is no doubt that mechanistic-based PK models have had a great impact on drug discovery and development in many different areas (Table 15.2),[80] such as bioavailability prediction, food effect prediction, formulation optimization, impact of drug substance properties on bioavailability, in vivo dissolution, GI metabolism, in vivo CL, first-pass metabolism prediction, cross-species extrapolation, and DDIs. As a result, regulatory agencies started to receive more and more submissions containing PBPK modeling.

The PBPK modeling approach has especially facilitated clinical pharmacology-related decisions.[81] In Zhao et al., PBPK modeling cases submitted for investigational new drugs (INDs) and new drug applications (NDAs) between Jul. 2008 and Jun. 2010 were summarized, which have facilitated the following types of regulatory decision making: the need to conduct specific clinical pharmacology studies, specific study designs, and appropriate labeling language.[81] In that paper, all the PBPK models were applied to address metabolism-related questions, such as DDI prediction; prediction of PK in hepatic impairment populations in extensive, intermediate, or poor metabolizers; and in pediatric populations.

Among the areas that PBPK modeling has been submitted in new drug application and evaluation, DDI prediction has been the main area of interest (Table 15.3). There are also examples where PBPK models were applied to predict PK in different populations: the levonorgestrel-releasing intrauterine system, and simeprevir oral capsules, and for absorption-related process prediction, such as predicting the p-glycoprotein (p-gp) contribution to oral absorption (ceritinib) and predicting the impact of stomach pH on absorption (panobinostat). Similar trends have been observed by the EMA, where DDI prediction is the major area of PBPK modeling applications and then PK in special populations.[82]

In 2015, a series of articles were published that summarized the experiences and viewpoints of experts in industry, European regulatory agency, and the US FDA regarding the application of PBPK modeling in drug development as a regulatory review.[80,82,83] Overall, all the parties agreed that PBPK modeling is a tool with great potential to assess benefits and risks. There are some areas in which scientists have high confidence, such as PK prediction for CYP-cleared substrates, absorption prediction for BCS I drugs, and DDI prediction involving reversible CYP inhibition alone or CYP induction alone.[80] Confidence levels for many other areas are still low, such as CL and absorption involving active transport; DDI prediction involving combined reversible, time-dependent inhibition; induction of CYPs; and absorption for BCS III and IV drugs.[80]

15.4.2 In generic drug evaluation

Unlike with NDAs, where firms have started submitting PBPK modeling in their application packages, PBPK modeling in generic drug applications is rare. The Office of Generic Drugs of the US FDA had

TABLE 15.3 New Drug Applications That Involved Physiologically Based PK Modeling (2009–15)

Drug name	Approval year	Area
Sildenafil[84,85]	2009	DDI
Cabazitaxel[86]	2010	In vivo intrinsic clearance prediction
Rivaroxaban[87,88]	2011	DDI
Vilazadone[89]	2011	DDI
Rilpivirine[90,91]	2011	DDI
Crizotinib[92]	2011	DDI
SKYLA (Levonorgestrol IUD)[93,94]	2013	PK among different populations
Bosutinib[95]	2012	DDI
Perampanel[96]	2012	DDI
Ponatinib[97,98]	2012	DDI
Canagliflozin[99]	2013	DDI
Simerprevir[100,101]	2013	DDI, PK in subpopulations
Ibrutinib (PCI32765)[102,103]	2013	DDI, PK in specific populations, GI luminal concentration
Macitentan[104,105]	2013	DDI
Naloxegol[106,107]	2014	DDI
Eliglustat[108,109]	2014	DDI
Ruxolitinib[110]	2014	DDI
Ceritinib[111,112]	2014	DDI, P-gp contribution to oral absorption
Olaparib[113,114]	2014	DDI
Blinatumomab (AMG103)[115,116]	2014	DDI
Rilpivirine[90,91]	2014	DDI
Belinostat[117]	2014	DDI
Lenvatinib[118,119]	2015	DDI
Panobinostat[120,121]	2015	DDI, impact of stomach pH on absorption

internally started exploring this type of predictive tools to address various issues in bioequivalence (BE) guidance development, citizen petition responses, abbreviated new drug application (ANDA) review consultations, protocol development for regulatory research studies, and investigation of postapproval issues.[1] In terms of the distribution of the number of projects using PBPK modeling in each area, the majority of the modeling activity has been in the BE guidance development and consult response (ie, internal needs to address questions).[122] Specifically, the PBPK absorption models have helped answer a number of questions, such as whether the in vivo dissolution profile correlates with the in vitro dissolution profile; whether the partial AUCs are appropriate as BE metrics; whether there will be in vivo alcohol dose dumping for specific formulations based on in vitro alcohol dose dumping studies; whether the GI local concentration/exposure correlates with the plasma concentration/exposure; the sensitivity of PK parameters to the change of formulation parameters, such as particle size; and the formulation parameters that affect PK BE conclusions.[122]

The main difference between the generic product and the reference drug product is the formulation difference. Mechanistic-based absorption and PK models that capture the formulation characteristics would be used to identify critical quality attributes by parameter sensitivity analysis and potentially very useful in product development. The challenge is verification of the prediction and building confidence in using this type of modeling and simulation tools to aid generic drug development and review.

15.5 SUMMARY

In this chapter, we first briefly introduced two basic modeling and simulation approaches for biopharmaceutical and PK: population PK-pharmacodynamic modeling and physiologically based PK modeling. Then we focused on introducing the theories of using mechanisms based on PBPK models to predict ADME. A search in PubMed was conducted to summarize the published case examples and areas where mechanistic PBPK models have been applied (Table 15.2). Finally, cases were introduced where mechanistic PBPK modeling and simulation were used in the new drug evaluation (Table 15.3), as well as generic drug evaluation.

Overall, mechanistic PBPK modeling and simulation have great potential for drug development and evaluation. The confidence level differs depending on the parameters to be predicted. There are also many scientific challenges connected with the development of better predictive models.

References

1. Zhang X, Lionberger RA. Modeling and simulation of biopharmaceutical performance. *Clin Pharmacol Ther* 2014;**95**(5):480–2.
2. Benet LZ. The role of BCS (biopharmaceutics classification system) and BDDCS (biopharmaceutics drug disposition classification system) in drug development. *J Pharm Sci* 2013;**102**(1):34–42.
3. Sheiner LB, Beal SL. Estimation of pooled pharmacokinetic parameters describing populations. In: Endrenyi L, editor. *Kinetic data analysis*. New York, NY: Plenum Press; 1981. p. 271–84.

4. Kiang TK, Sherwin CM, Spigarelli MG, Ensom MH. Fundamentals of population pharmacokinetic modelling: modelling and software. *Clin Pharmacokinet* 2012;**51**(8):515–25.
5. Sheiner LB. Learning versus confirming in clinical drug development. *Clin Pharmacol Ther* 1997;**61**(3):275–91.
6. FDA. Guidance for industry population pharmacokinetics. ed; 1999.
7. EMA. Guideline on reporting the results of population pharmacokinetic analysis. ed; 2006.
8. Jamei M, Dickinson GL, Rostami-Hodjegan A. A framework for assessing inter-individual variability in pharmacokinetics using virtual human populations and integrating general knowledge of physical chemistry, biology, anatomy, physiology and genetics: a tale of "bottom-up" vs "top-down" recognition of covariates. *Drug Metab Pharmacokinet* 2009;**24**(1):53–75.
9. Rowland M, Peck C, Tucker G. Physiologically-based pharmacokinetics in drug development and regulatory science. *Annu Rev Pharmacol Toxicol* 2011;**51**:45–73.
10. Jones H, Rowland-Yeo K. Basic concepts in physiologically based pharmacokinetic modeling in drug discovery and development. *CPT: Pharmacometrics Syst Pharmacol* 2013;**2**:e63.
11. Thompson MD, Beard DA. Physiologically based pharmacokinetic tissue compartment model selection in drug development and risk assessment. *J Pharm Sci* 2012;**101**(1):424–35.
12. Jones HM, Parrott N, Jorga K, Lave T. A novel strategy for physiologically based predictions of human pharmacokinetics. *Clin Pharmacokinet* 2006;**45**(5):511–42.
13. Thompson MD, Beard DA. Development of appropriate equations for physiologically based pharmacokinetic modeling of permeability-limited and flow-limited transport. *J Pharmacokinet Pharmacodyn* 2011;**38**(4):405–21.
14. Khalil F, Laer S. Physiologically based pharmacokinetic modeling: methodology, applications, and limitations with a focus on its role in pediatric drug development. *J Biomed Biotechnol* 2011;**2011**:907461.
15. Jones HM, Gardner IB, Collard WT, Stanley PJ, Oxley P, Hosea NA, et al. Simulation of human intravenous and oral pharmacokinetics of 21 diverse compounds using physiologically based pharmacokinetic modelling. *Clin Pharmacokinet* 2011;**50**(5):331–47.
16. Price PS, Conolly RB, Chaisson CF, Gross EA, Young JS, Mathis ET, et al. Modeling interindividual variation in physiological factors used in PBPK models of humans. *Crit Rev Toxicol* 2003;**33**(5):469–503.
17. Willmann S, Hohn K, Edginton A, Sevestre M, Solodenko J, Weiss W, et al. Development of a physiology-based whole-body population model for assessing the influence of individual variability on the pharmacokinetics of drugs. *J Pharmacokinet Pharmacodyn* 2007;**34**(3):401–31.
18. Kostewicz ES, Aarons L, Bergstrand M, Bolger MB, Galetin A, Hatley O, et al. PBPK models for the prediction of in vivo performance of oral dosage forms. *Eur J Pharm Sci* 2014;**57**:300–21.
19. FDA. Guidance for industry: drug interaction studies—study design, data analysis, implications for dosing, and labeling recommendations. ed; 2012.
20. EMA. Concept paper on qualification and reporting of physiologically-based pharmacokinetic (PBPK) modelling and analyses. ed., United Kingdom; 2014.
21. Zhou R, Moench P, Heran C, Lu X, Mathias N, Faria TN, et al. pH-dependent dissolution in vitro and absorption in vivo of weakly basic drugs: development of a canine model. *Pharm Res* 2005;**22**(2):188–92.
22. Cao X, Yu LX, Sun D. Drug absorption principles. In: Krishna R, Yu LX, editors. Biopharmaceutics applications in drug development. New York, NY: Springer; 2008. p. 75–100.
23. Hattori Y, Haruna Y, Otsuka M. Dissolution process analysis using model-free Noyes–Whitney integral equation. *Colloids and Surf B Biointerfaces* 2013;**102**:227–31.
24. Yu LX, Amidon GL. A compartmental absorption and transit model for estimating oral drug absorption. *Int J Pharm* 1999;**186**(2):119–25.
25. Huang W, Lee SL, Yu LX. Mechanistic approaches to predicting oral drug absorption. *AAPS J* 2009;**11**(2):217–24.
26. Cao X, Gibbs ST, Fang L, Miller HA, Landowski CP, Shin HC, et al. Why is it challenging to predict intestinal drug absorption and oral bioavailability in human using rat model. *Pharm Res* 2006;**23**(8):1675–86.
27. Lennernas H. Human in vivo regional intestinal permeability: importance for pharmaceutical drug development. *Mol Pharm* 2014;**11**(1):12–23.
28. Salphati L, Childers K, Pan L, Tsutsui K, Takahashi L. Evaluation of a single-pass intestinal-perfusion method in rat for the prediction of absorption in man. *J Pharm Pharmacol* 2001;**53**(7):1007–13.
29. Jamei M, Turner D, Yang J, Neuhoff S, Polak S, Rostami-Hodjegan A, et al. Population-based mechanistic prediction of oral drug absorption. *AAPS J* 2009;**11**(2):225–37.
30. Freitas AA, Limbu K, Ghafourian T. Predicting volume of distribution with decision tree-based regression methods using predicted tissue:plasma partition coefficients. *J Cheminform* 2015;**7**:6.
31. Poulin P, Theil FP. Prediction of pharmacokinetics prior to in vivo studies. 1. Mechanism-based prediction of volume of distribution. *J Pharm Sci* 2002;**91**(1):129–56.
32. Berezhkovskiy LM. Determination of volume of distribution at steady state with complete consideration of the kinetics of protein and tissue binding in linear pharmacokinetics. *J Pharm Sci* 2004;**93**(2):364–74.
33. Rodgers T, Rowland M. Physiologically based pharmacokinetic modelling 2: predicting the tissue distribution of acids, very weak bases, neutrals and zwitterions. *J Pharm Sci* 2006;**95**(6):1238–57.
34. Rodgers T, Leahy D, Rowland M. Physiologically based pharmacokinetic modeling 1: predicting the tissue distribution of moderate-to-strong bases. *J Pharm Sci* 2005;**94**(6):1259–76.
35. Westerhout J, Ploeger B, Smeets J, Danhof M, de Lange EC. Physiologically based pharmacokinetic modeling to investigate regional brain distribution kinetics in rats. *AAPS J* 2012;**14**(3):543–53.
36. Komura H, Iwaki M. In vitro and in vivo small intestinal metabolism of CYP3A and UGT substrates in preclinical animals species and humans: species differences. *Drug Metab Rev* 2011;**43**(4):476–98.
37. Paine MF, Hart HL, Ludington SS, Haining RL, Rettie AE, Zeldin DC. The human intestinal cytochrome P450 "pie". *Drug Metab Dispos* 2006;**34**(5):880–6.
38. Darwich AS, Neuhoff S, Jamei M, Rostami-Hodjegan A. Interplay of metabolism and transport in determining oral drug absorption and gut wall metabolism: a simulation assessment using the "Advanced Dissolution, Absorption, Metabolism (ADAM)" model. *Curr Drug Metab* 2010;**11**(9):716–29.
39. Guengerich FP. A malleable catalyst dominates the metabolism of drugs. *Proc Natl Acad Sci USA* 2006;**103**(37):13565–6.
40. Yang J, Jamei M, Yeo KR, Tucker GT, Rostami-Hodjegan A. Prediction of intestinal first-pass drug metabolism. *Curr Drug Metab* 2007;**8**(7):676–84.
41. Karlsson FH, Bouchene S, Hilgendorf C, Dolgos H, Peters SA. Utility of in vitro systems and preclinical data for the prediction of human intestinal first-pass metabolism during drug discovery and preclinical development. *Drug Metab Dispos* 2013;**41**(12):2033–46.
42. Gertz M, Harrison A, Houston JB, Galetin A. Prediction of human intestinal first-pass metabolism of 25 CYP3A substrates from in vitro clearance and permeability data. *Drug Metab Dispos* 2010;**38**(7):1147–58.
43. Heikkinen AT, Baneyx G, Caruso A, Parrott N. Application of PBPK modeling to predict human intestinal metabolism of CYP3A substrates—an evaluation and case study using GastroPlus. *Eur J Pharm Sci* 2012;**47**(2):375–86.

44. Zhao Y, Hebert MF, Venkataramanan R. Basic obstetric pharmacology. *Semin Perinatol* 2014;**38**(8):475–86.
45. Liu L, Pang KS. An integrated approach to model hepatic drug clearance. *Eur J Pharm Sci* 2006;**29**(3-4):215–30.
46. Chao P, Uss AS, Cheng KC. Use of intrinsic clearance for prediction of human hepatic clearance. *Expert Opin Drug Metab Toxicol* 2010;**6**(2):189–98.
47. Paixao P, Gouveia LF, Morais JA. Prediction of the human oral bioavailability by using in vitro and in silico drug related parameters in a physiologically based absorption model. *Int J Pharm* 2012;**429**(1-2):84–98.
48. Nakamaru Y, Emoto C, Shimizu M, Yamazaki H. Human pharmacokinetic profiling of the dipeptidyl peptidase-IV inhibitor teneligliptin using physiologically based pharmacokinetic modeling. *Biopharm Drug Dispos* 2015;**36**(3):148–62.
49. Poirier A, Cascais AC, Funk C, Lave T. Prediction of pharmacokinetic profile of valsartan in human based on in vitro uptake transport data. *J Pharmacokinet Pharmacodyn* 2009;**36**(6):585–611.
50. De Buck SS, Sinha VK, Fenu LA, Nijsen MJ, Mackie CE, Gilissen RA. Prediction of human pharmacokinetics using physiologically based modeling: a retrospective analysis of 26 clinically tested drugs. *Drug Metab Dispos* 2007;**35**(10):1766–80.
51. Lu C, Suri A, Shyu WC, Prakash S. Assessment of cytochrome P450-mediated drug–drug interaction potential of orteronel and exposure changes in patients with renal impairment using physiologically based pharmacokinetic modeling and simulation. *Biopharm Drug Dispos* 2014;**35**(9):543–52.
52. Thiel C, Schneckener S, Krauss M, Ghallab A, Hofmann U, Kanacher T, et al. A systematic evaluation of the use of physiologically based pharmacokinetic modeling for cross-species extrapolation. *J Pharm Sci* 2015;**104**(1):191–206.
53. Xia B, Heimbach T, Lin TH, Li S, Zhang H, Sheng J, et al. Utility of physiologically based modeling and preclinical in vitro/in vivo data to mitigate positive food effect in a BCS class 2 compound. *AAPS PharmSciTech* 2013;**14**(3):1255–66.
54. Heimbach T, Xia B, Lin TH, He H. Case studies for practical food effect assessments across BCS/BDDCS class compounds using in silico, in vitro, and preclinical in vivo data. *AAPS J* 2013;**15**(1):143–58.
55. Zhang H, Xia B, Sheng J, Heimbach T, Lin TH, He H, et al. Application of physiologically based absorption modeling to formulation development of a low solubility, low permeability weak base: mechanistic investigation of food effect. *AAPS PharmSciTech* 2014;**15**(2):400–6.
56. Patel N, Polak S, Jamei M, Rostami-Hodjegan A, Turner DB. Quantitative prediction of formulation-specific food effects and their population variability from in vitro data with the physiologically-based ADAM model: a case study using the BCS/BDDCS Class II drug nifedipine. *Eur J Pharm Sci* 2014;**57**:240–9.
57. Wagner C, Thelen K, Willmann S, Selen A, Dressman JB. Utilizing in vitro and PBPK tools to link ADME characteristics to plasma profiles: case example nifedipine immediate release formulation. *J Pharm Sci* 2013;**102**(9):3205–19.
58. Shono Y, Jantratid E, Kesisoglou F, Reppas C, Dressman JB. Forecasting in vivo oral absorption and food effect of micronized and nanosized aprepitant formulations in humans. *Eur J Pharm Biopharm* 2010;**76**(1):95–104.
59. Wagner C, Jantratid E, Kesisoglou F, Vertzoni M, Reppas C, Dressman JB. Predicting the oral absorption of a poorly soluble, poorly permeable weak base using biorelevant dissolution and transfer model tests coupled with a physiologically based pharmacokinetic model. *Eur J Pharm Biopharm* 2012;**82**(1):127–38.
60. Shono Y, Jantratid E, Janssen N, Kesisoglou F, Mao Y, Vertzoni M, et al. Prediction of food effects on the absorption of celecoxib based on biorelevant dissolution testing coupled with physiologically based pharmacokinetic modeling. *Eur J Pharm Biopharm* 2009;**73**(1):107–14.
61. Wang Y, Wang Z, Zuo Z, Tomlinson B, Lee BT, Bolger MB, et al. Clinical pharmacokinetics of buffered propranolol sublingual tablet (Promptol)-application of a new "physiologically based" model to assess absorption and disposition. *AAPS J* 2013;**15**(3):787–96.
62. Wong H, Theil FP, Cui Y, Marsters Jr. JC, Khojasteh SC, Vernillet L, et al. Interplay of dissolution, solubility, and nonsink permeation determines the oral absorption of the Hedgehog pathway inhibitor GDC-0449 in dogs: an investigation using preclinical studies and physiologically based pharmacokinetic modeling. *Drug Metab Dispos* 2010;**38**(7):1029–38.
63. Chiang PC, La H, Zhang H, Wong H. Systemic concentrations can limit the oral absorption of poorly soluble drugs: an investigation of non-sink permeation using physiologically based pharmacokinetic modeling. *Mol Pharm* 2013;**10**(11):3980–8.
64. Gao Y, Carr RA, Spence JK, Wang WW, Turner TM, Lipari JM, et al. A pH-dilution method for estimation of biorelevant drug solubility along the gastrointestinal tract: application to physiologically based pharmacokinetic modeling. *Mol Pharm* 2010;**7**(5):1516–26.
65. Chiang PC, Wong H. Incorporation of physiologically based pharmacokinetic modeling in the evaluation of solubility requirements for the salt selection process: a case study using phenytoin. *AAPS J* 2013;**15**(4):1109–18.
66. Chetty M, Rose RH, Abduljalil K, Patel N, Lu G, Cain T, et al. Applications of linking PBPK and PD models to predict the impact of genotypic variability, formulation differences, differences in target binding capacity and target site drug concentrations on drug responses and variability. *Front Pharmacol* 2014;**5**:258.
67. Dong D, Wang X, Wang H, Zhang X, Wang Y, Wu B. Elucidating the in vivo fate of nanocrystals using a physiologically based pharmacokinetic model: a case study with the anticancer agent SNX-2112. *Int J Nanomedicine* 2015;**10**:2521–35.
68. Olivares-Morales A, Kamiyama Y, Darwich AS, Aarons L, Rostami-Hodjegan A. Analysis of the impact of controlled release formulations on oral drug absorption, gut wall metabolism and relative bioavailability of CYP3A substrates using a physiologically-based pharmacokinetic model. *Eur J Pharm Sci* 2015;**67**:32–44.
69. Berlin M, Ruff A, Kesisoglou F, Xu W, Wang MH, Dressman JB. Advances and challenges in PBPK modeling—analysis of factors contributing to the oral absorption of atazanavir, a poorly soluble weak base. *Eur J Pharm Biopharm* 2015;**93**:267–80.
70. Otsuka K, Wagner C, Selen A, Dressman J. Prediction of in-vivo pharmacokinetic profile for immediate and modified release oral dosage forms of furosemide using an in-vitro-in-silico-in-vivo approach. *J Pharm Pharmacol* 2015;**67**(5):651–65.
71. Kambayashi A, Blume H, Dressman J. Understanding the in vivo performance of enteric coated tablets using an in vitro-in silico-in vivo approach: case example diclofenac. *Eur J Pharm Biopharm* 2013;**85**(3 Pt B):1337–47.
72. Sjogren E, Westergren J, Grant I, Hanisch G, Lindfors L, Lennernas H, et al. In silico predictions of gastrointestinal drug absorption in pharmaceutical product development: application of the mechanistic absorption model GI-Sim. *Eur J Pharm Sci* 2013;**49**(4):679–98.
73. Willmann S, Edginton AN, Kleine-Besten M, Jantratid E, Thelen K, Dressman JB. Whole-body physiologically based pharmacokinetic population modelling of oral drug administration: inter-individual variability of cimetidine absorption. *J Pharm Pharmacol* 2009;**61**(7):891–9.
74. Willmann S, Thelen K, Becker C, Dressman JB, Lippert J. Mechanism-based prediction of particle size-dependent dissolution and absorption: cilostazol pharmacokinetics in dogs. *Eur J Pharm Biopharm* 2010;**76**(1):83–94.
75. Cai H, Stoner C, Reddy A, Freiwald S, Smith D, Winters R, et al. Evaluation of an integrated in vitro-in silico PBPK (physiologically based pharmacokinetic) model to provide estimates of human bioavailability. *Int J Pharm* 2006;**308**(1-2):133–9.

76. Peng HT, Bouak F, Vartanian O, Cheung B. A physiologically based pharmacokinetics model for melatonin--effects of light and routes of administration. *Int J Pharm* 2013;**458**(1):156–68.
77. Johnson TN, Zhou D, Bui KH. Development of physiologically based pharmacokinetic model to evaluate the relative systemic exposure to quetiapine after administration of IR and XR formulations to adults, children and adolescents. *Biopharm Drug Dispos* 2014;**35**(6):341–52.
78. Kambayashi A, Blume H, Dressman JB. Predicting the oral pharmacokinetic profiles of multiple-unit (pellet) dosage forms using a modeling and simulation approach coupled with biorelevant dissolution testing: case example diclofenac sodium. *Eur J Pharm Biopharm* 2014;**87**(2):236–43.
79. Peters SA. Identification of intestinal loss of a drug through physiologically based pharmacokinetic simulation of plasma concentration-time profiles. *Clin Pharmacokinet* 2008;**47**(4):245–59.
80. Jones HM, Chen Y, Gibson C, Heimbach T, Parrott N, Peters SA, et al. Physiologically based pharmacokinetic modeling in drug discovery and development: a pharmaceutical industry perspective. *Clin Pharmacol Ther* 2015;**97**(3):247–62.
81. Zhao P, Zhang L, Grillo JA, Liu Q, Bullock JM, Moon YJ, et al. Applications of physiologically based pharmacokinetic (PBPK) modeling and simulation during regulatory review. *Clin Pharmacol Ther* 2011;**89**(2):259–67.
82. Shepard T, Scott G, Cole S, Nordmark A, Bouzom F. Physiologically based models in regulatory submissions: output from the ABPI/MHRA forum on physiologically based modeling and simulation. *Clin Pharmacol Ther* 2015;**4**(4):221–5.
83. Wagner C, Zhao P, Pan Y, Hsu V, Grillo J, Huang SM, et al. Application of Physiologically based pharmacokinetic (PBPK) modeling to support dose selection: report of an FDA public workshop on PBPK. *Clin Pharmacol Ther* 2015;**4**(4):226–30.
84. Sildenafil review package. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2009/022473s000_ClinPharmR.pdf.
85. Sildenafil label. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/label/2012/022473s003lbl.pdf.
86. Cabazitaxel review package. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2010/201023s000ClinPharmR.pdf.
87. Rivaroxaban review package. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/022406Orig1s000ClinPharmR.pdf.
88. Rivaroxaban label. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/label/2011/022406s000lbl.pdf.
89. Vilazadone review package [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/022567Orig1s000ClinPharmR.pdf.
90. Rilpivirine review package. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/202022Orig1s000ClinPharmR.pdf.
91. Rilpivirine label. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/label/2011/202022s000lbl.pdf.
92. Crizotinib review package. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2011/202570Orig1s000Admincorres_R.pdf.
93. Levonorgestrol IUD review package. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2013/203159Orig1s000ClinPharmR.pdf.
94. Levonorgestrol IUD label. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/label/2013/203159s002lbledt1.pdf.
95. Bosutinib review package. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2012/203341Orig1s000ClinPharmR.pdf.
96. Perampanel review package. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2012/202834Orig1s000AdminCorres.pdf.
97. Ponatinib review package. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2012/203469Orig1s000ClinPharmR.pdf.
98. Ponatinib label. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/label/2012/203469lbl.pdf.
99. Canagliflozin review package. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2013/204042Orig1s000ClinPharmR.pdf.
100. Simerprevir review package. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2013/205123Orig1s000ClinPharmR.pdf.
101. Simerprevir label. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/label/2013/205123s001lbl.pdf.
102. Ibrutinib review package. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2013/205552Orig1s000ClinPharmR.pdf
103. Ibrutinib label [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/label/2013/205552s000lbl.pdf.
104. Macitentan review package. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2013/204410Orig1s000ClinPharmR.pdf.
105. Macitentan label. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/label/2013/204410s000lbl.pdf.
106. Naloxegol review package. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2014/204760Orig1s000ClinPharm.pdf.
107. Naloxegol label. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/label/2014/204760s000lbl.pdf.
108. Eliglustat review package. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2014/205494Orig1s000ClinPharmR.pdf.
109. Eliglustat label. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/label/2014/205494Orig1s000lbl.pdf.
110. Ruxolitinib label. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/label/2014/202192s006lbl.pdf.
111. Ceritinib review package. 2014 [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2014/205755Orig1s000ClinPharmR.pdf.
112. Ceritinib label. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/label/2014/205755s000lbl.pdf.
113. Olaparib review package. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2014/206162Orig1s000ClinPharmR.pdf.
114. Olaparib label. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/label/2014/206162lbl.pdf.
115. Blinatumomab review package. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2014/125557Orig1s000ClinPharmR.pdf
116. Blinatumomab label. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/label/2014/125557lbl.pdf.

117. Belinostat review package. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2014/206256Orig1s000ClinPharmR.pdf
118. Lenvatinib review package. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2015/206947Orig1s000ClinPharmR.pdf
119. Lenvatinib label. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/label/2015/206947s000lbl.pdf.
120. Panobinostat review package [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2015/205353Orig1s000ClinPharmR.pdf.
121. Panobinostat label. [05/30/2015)]; Available from: http://www.accessdata.fda.gov/drugsatfda_docs/label/2015/205353s000lbl.pdf.
122. Lionberger R. Application of PBPK models in assessment of bioequivalence. From: https://zerista.s3.amazonaws.com/item_files/0df5/attachments/31736/original/248.pdf; 2014 [retrieved 30.05.15].

This page intentionally left blank

CHAPTER

16

In Vitro/In Vivo Correlations: Fundamentals, Development Considerations, and Applications

Y. Qiu[1] *and J.Z. Duan*[2]

[1]AbbVie, Inc., North Chicago, IL, United States [2]U.S. Food and Drug Administration, Silver Spring, MD, United States

16.1 INTRODUCTION

For decades, developing in vitro tests and models to assess or predict the in vivo performance of pharmaceutical products has been sought after as a means of screening, optimizing, and monitoring dosage forms. With solid oral dosage forms, it most frequently starts with an attempt to link the results of an in vitro release test to in vivo pharmacokinetic studies. Through exploring the association or relationship between in vitro dissolution/release and in vivo absorption data, an in vitro/in vivo relationship (IVIVR) may be identified for a drug product. Such a relationship is often qualitative or semiquantitative in nature (eg, rank order). When a predictive relationship or model is established and validated between in vitro dissolution and in vivo absorption profiles, it is designated as in vitro/in vivo correlation (IVIVC).

IVIVC of oral solid products have received considerable attention from the industry, regulatory agencies, and academia over the past two decades, particularly since the publication of Food and Drug Administration (FDA) guidance of dissolution testing of immediate-release (IR) and IVIVC of extended-release (ER) dosage forms in 1997[1,2] and subsequent guidelines issued by European regulatory authorities (EMA).[3–5] As a result, there has been increased confidence, effort, and success using in vitro tests to evaluate or predict in vivo performance of solid drug products, especially ER dosage forms, based on IVIVC.[6–13] With an established IVIVC, the dissolution data can be used not only as a quality control tool but also for guiding and optimizing product development, setting meaningful specifications, and serving as a surrogate for a bioavailability study. This chapter will discuss basic principles and methodology utilized to establish and evaluating IVIVC models, as well as applications of IVIVC in developing solid dosage forms with a primary focus on ER products. Additional topics include the importance of understanding drug substance, dosage form, their in vitro and in vivo behaviors, and in vitro test methods to explore and develop IVIVC.

16.1.1 In vitro/in vivo correlation

IVIVC is defined by United States Pharmacopeia (USP) and the FDA respectively as follows[2,14]:

> *USP:* the establishment of a relationship between a biological property, or a parameter derived from a biological property produced by a dosage form, and a physicochemical characteristic of the same dosage form.
> *FDA:* a predictive mathematical model describing the relationship between an in vitro property (usually the extent or rate of drug release) and a relevant in vivo response (eg, plasma concentration or amount of drug absorbed).

Evaluation of IVIVCs by different levels was first proposed for oral dosage forms in the USP's information chapter <1088>[15] and was later adopted globally. Presently, IVIVC is categorized by the FDA into levels A, B, C, and Multiple C depending upon the type of data used to establish the relationship and ability of the correlation to predict the complete plasma profile of a dosage form.[2]

DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00016-9

Level A: A predictive mathematical model for the relationship between the entire in vitro release time course and the entire in vivo response time course, for example, the time course of plasma drug concentration or amount of drug absorbed.
Level B: A predictive mathematical model for the relationship between summary parameters that characterize the in vitro and in vivo time courses, for example, models that relate the mean in vitro dissolution time to the mean in vivo dissolution time, or to mean residence time (MRT) in vivo.
Level C: A predictive mathematical model for the relationship between the amount dissolved in vitro at a particular time (eg, Q_{60}) or the time required for dissolution of a fixed amount (eg, $T_{50\%}$) and a summary parameter that characterizes the in vivo time course (eg, C_{max} or AUC).
Multiple Level C: Predictive mathematical models for the relationships between the amount dissolved at several time points of the product and one or several pharmacokinetic parameters of interest.

Level A is the most informative and most useful from both scientific and regulatory perspectives in that it represents a point-to-point relationship between in vitro release and in vivo release/absorption from the dosage form. It can be used to predict the entire in vivo time course from the in vitro data. Multiple Level C is also useful as it provides the in vitro release profile of a dosage form with biological meaning. Level C can be useful in early stages of product development or setting meaningful specifications, although it does not reflect the complete shape of the plasma concentration-time curve. Level B utilizes the principles of statistical moment analysis. However, it is the least useful for regulatory applications because different in vitro or in vivo profiles may produce similar mean time values.

16.1.2 IVIVC and product development

The value of IVIVC in product development has been recognized since the early 1960s. Exploring in vitro/in vivo association or correlation is very useful in guiding formulation and process development. A validated IVIVC can support formulation and process changes and scale-up, help develop meaningful dissolution specifications, and support the use of dissolution as a surrogate for an in vivo study, since IVIVC provides a biological meaning to the results of the in vitro test.[2] Thus, availability of a validated predictive IVIVC can result in a significant positive impact on product quality, development efficiency, and reduced regulatory burden.

The approaches and challenges of developing an IVIVC have undergone extensive discussion and debate since the 1980s.[16–25] In general, there is increased uncertainty associated with developing an IVIVC for IR oral dosage forms because the in vivo apparent drug absorption is often a function of a multitude of variables, many of which are difficult to isolate or mimic in vitro. For example, a correlation between the in vitro dissolution and the in vivo absorption for IR dosage forms of highly water soluble drugs (eg, Biopharmaceutical Classification System 1 and 3) is usually not possible because gastric emptying or membrane permeation is often the rate-limiting step. Absorptive and efflux transporters and/or gut metabolism can play a significant role in the apparent absorption of BCS 2–4 compounds[26–28] depending on the dose, physicochemical, biopharmaceutical properties, and dissolution rate, making correlating dissolution with absorption difficult.[29] Compared to IR products, an IVIVC is more suitable for ER dosage forms where drug release is, by design, rate limiting in the absorption process. In addition, because a patient is typically exposed to a specific range of plasma levels over an extended period of time (eg, up to 24 hours) following administration of a modified-release (MR) dosage form, an in vitro test method with a qualitative or quantitative relationship with in vivo data is desired to assure the consistent in vivo performance. Hence, the FDA has recommended investigation of the possibility of an IVIVC in the development of ER dosage forms.[2]

16.2 DEVELOPMENT AND ASSESSMENT OF AN IVIVC

The regulatory guidance on IVIVC of ER oral dosage forms issued by the FDA in 1997 provides a comprehensive scientific framework and regulatory guideline to IVIVC model development, evaluation, and applications. In general, establishing an IVIVC consists of (1) study design, (2) model building, and (3) model validation based on an appropriate statistical assessment.

16.2.1 Study design and general considerations

Development of an IVIVC requires in vitro and in vivo data of formulations with varying in vitro release rates and corresponding in vivo differences. These data may come from studies at the early or late stage of product development, such as bioavailability studies conducted in the formulation screening stage or an in vivo study specifically designed to explore IVIVC.[11,30,31] The in vitro release rates, as measured by percent dissolved for each formulation studied, should differ adequately (eg, by 10%).[2]

To obtain useful data for IVIVC, discriminating dissolution methodology is essential. Based on the FDA

guidance,[2] in vitro data are preferably generated in an aqueous medium using USP apparatus I (basket) or II (paddle), operating within an appropriate range of rotation speeds (eg, 50–100 rpm). In other cases, USP apparatus III (reciprocating cylinder) or IV (flow-through cell) may also be used. Generally, any in vitro test method may be used to obtain the dissolution characteristics of the dosage forms as long as it is shown to be predictive of the in vivo performance. The dissolution profiles of at least 12 individual dosage units should be determined for IVIVC purposes. The coefficient of variation (CV) for the mean dissolution profiles of a single batch should be less than 20% at early time points and less than 10% at other time points.

According to the FDA guidance,[2] bioavailability studies for IVIVC development should be performed in humans with enough subjects in order to adequately characterize the absorption profiles of the drug product. Although crossover studies are preferred, parallel studies or cross-study analyses are also acceptable. The latter may involve normalization with a common reference treatment. The reference product in developing an IVIVC may be an intravenous solution, an aqueous oral solution, or an immediate-release product of the drug. In addition, IVIVCs are usually developed in the fasted state. When a drug is not tolerated in the fasted state, studies may be conducted in the fed state.

16.2.2 IVIVC modeling

The principles and methodologies of IVIVC modeling and assessment have been extensively addressed and reviewed in the literature.[32–34] Developing an IVIVC model begins with understanding the following mathematical principles for characterizing in vivo drug release/absorption profiles or parameters associated with different types of IVIVC.

16.2.2.1 Convolution and deconvolution approaches used in Level A correlation

Convolution and deconvolution methods are essential tools for establishing Level A IVIVC. Convolution is a model-independent method based on linear system theory.

A linear system has the property that the response to a linear combination of inputs is the same linear combination of the individual responses (superposition). If the input is shifted in time by some amount, and the output is simply shifted by the same amount, such a system that is not sensitive to the time origin is known as a linear time-invariant (LTI) system.[35]

Once the impulse response function of a LTI system (ie, the way the system responds to a unit impulse) is measured, how the system will respond to any other possible inputs can be predicted in principle. This is because (1) the inputs to the system can be decomposed as a linear combination of some basic inputs, and (2) responses (or outputs) can be constructed as the same linear combination of the responses to each of the basic inputs in a wide variety of ways as illustrated in Fig. 16.1.

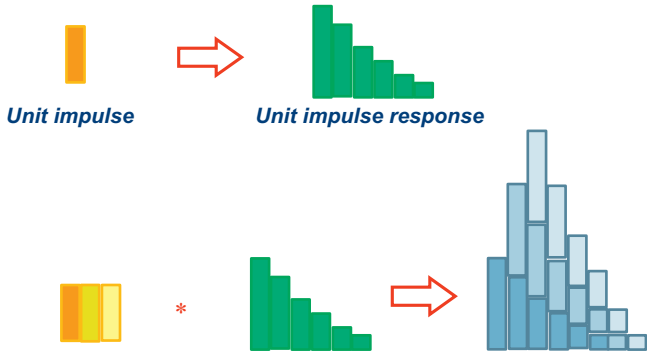


FIGURE 16.1 Illustration of input-response relationship of a LTI system (* denotes the convolution operation).

In pharmaceutical applications, drug disposition in the body is considered an LTI system. One of the measurable responses to the drug input (absorption) would be the plasma concentration-time profile. The response to an instantaneous (impulse) input would be the plasma concentration-time profile of following an intravenous bolus injection that is unique to individual drug molecules. Therefore, through applying the superposition principle of an LTI system in the study of the drug absorption process, a response, $C(t)$, to an arbitrary input, $f(t)$, of the system can be obtained using the following convolution integral[36]:

$$C(t) = f(t)^{*}C_{\delta}(t) = \int_{0}^{\infty} C_{\delta}(t-\tau)f(\tau)d\tau \qquad (16.1)$$

where $C_{\delta}(t)$ is the unit impulse response (UIR) that defines the characteristic of the system. It is the response of the system to an instantaneous unit input, usually attainable from an IV bolus or oral solution. By the same principle, $f(t)$ can be obtained by deconvolution, the inverse operation of convolution. Their applications in IVIVC are illustrated in Fig. 16.2, and representative systems are provided in Table 16.1. The definition of a system is flexible and is determined by the nature of the time functions involved.[37] Depending on the specific $C_{\delta}(t)$ and input responses used to define a system, $f(t)$ obtained by deconvolution in IVIVC may represent the dissolution process, absorption process, or the combined processes of the two.

In exploring IVIVC, Level A correlation is usually estimated by a two-stage procedure, that is, deconvolution

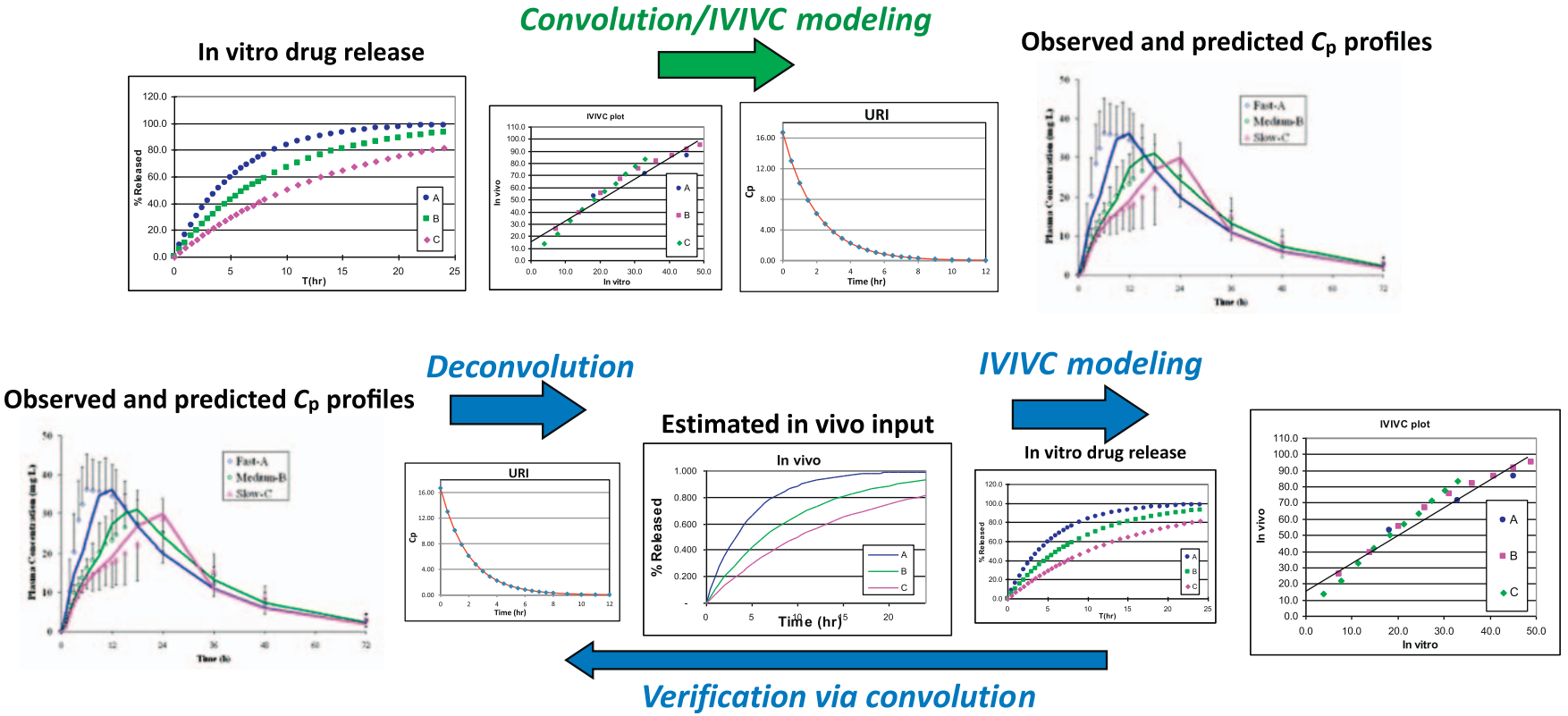
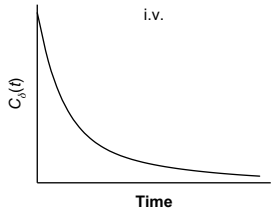
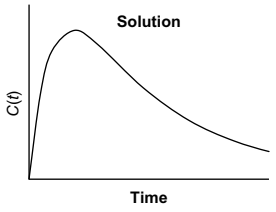
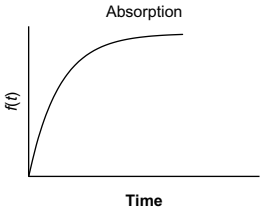
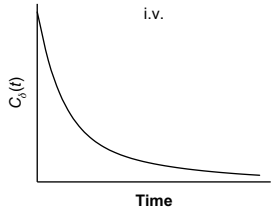
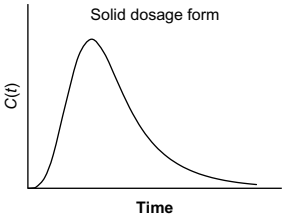
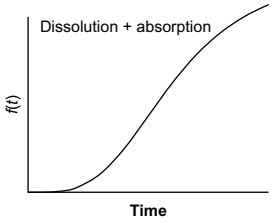
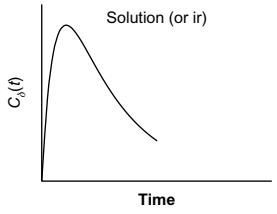
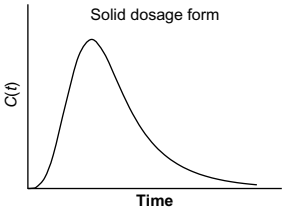
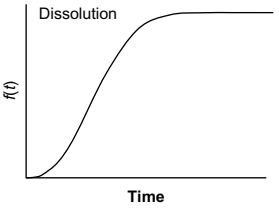


FIGURE 16.2 Illustration of convolution and deconvolution in IVIVC development.

TABLE 16.1 Illustration of System Definitions for Oral Administration

Case	Unit impulse response $C_\delta(t)$	Input response $C(t)$	Input function $f(t)$
I	Plasma level profile of IV bolus 	Plasma level profile of oral solution 	Absorption in the GIT 
II	Plasma level profile of IV bolus 	Plasma level profile of oral solid dosage form 	Dissolution and absorption in the GIT 
III	Plasma level profile of oral solution (or IR dosage form as an approximation) 	Plasma level profile of oral solid dosage form 	Dissolution in the GIT 

followed by correlating the fraction dissolved in vitro with the fraction released or absorbed in vivo. It may also be evaluated via a single-stage procedure, that is, a direct comparison of the observed with the predicted plasma concentration-time profiles obtained by convolution of the in vitro data and UIR. According to Eq. (16.1), the in vitro drug release and the in vivo input (release/absorption) estimated by deconvolution of the UIR with the observed plasma data are either directly superimposable or may be made to be superimposable by the use of a scaling factor when a 1:1 IVIVC exists. Similarly, the plasma concentration profile observed following oral administration should be in good agreement with that obtained by convolution of the UIR with the in vitro release data if there is a Level A IVIVC.

16.2.2.1.1 General solution

The exact solution of convolution or deconvolution can be obtained by operation of Laplace transform if each functional form is defined:

$$L\{C(t)\} = L\{(C_\delta {}^* f)(t)\} = L\{C_\delta(t)\}L\{f(t)\} \tag{16.2}$$

$$f(t) = L^{-1}\{\bar{f}(s)\} = L^{-1}\left\{\frac{\bar{c}(s)}{\bar{c}_\delta(s)}\right\} \tag{16.3}$$

where L and L^{-1} denote Laplace transform and inverse Laplace transform, respectively. Deconvolution methods include explicit (numerical point-area and midpoint methods, least squares curve fitting using polyexponential, polynomial, spline functions) and implicit methods (prescribed function or deconvolution via convolution).[38–50]

Since the disposition of most drugs can be described by polyexponentials,

$$C_\delta(t) = \sum_{i=1}^{n} A_i e^{-\alpha_i t} \tag{16.4}$$

the in vivo input function $f(t)$ can be obtained using Eq. (16.3). For example, in the case of single-exponential disposition ($n = 1$), $C_\delta(t) = A_1 e^{-\alpha_1 t}$, and hence, $f(t)$, the input rate, is given by [40]:

$$f(t) = \frac{[C'(t) + \alpha_1 C(t)]}{A_1} \tag{16.5}$$

The amount of drug absorbed from time 0 to t, $X_a(t)$, is then obtained by integration:

$$X_a(t) = \int_0^t f(t)dt = \frac{\left[C(t) + \alpha_1 \int_0^t C(t)dt\right]}{A_1} \qquad (16.6)$$

In cases where $C(t)$ or $f(t)$ cannot be fitted to an explicit function, numerical methods are used to deal with the raw data.

16.2.2.1.2 Numerical deconvolution

It is well recognized that classical linear compartmental kinetic models exhibit superposition linearity due to their origin in linear differential equations. Hence, linear systems analysis, especially in the form of numerical algorithms, is conceptually simple, and thus it is a useful tool for assessing absorption and IVIVC. Deconvolution using purely numerical algorithms has been used for evaluation of IVIVC since the 1960s.[51–53] The general process for deconvolution usually uses the basic principle of deconvolution through convolution (DTC) to determine the input function. The DTC method is an iterative procedure consisting of three steps. The input function is first defined by the selection of its initial parameter values followed by convolution with UIR function to calculate the response (drug concentrations). Subsequently, the agreement between the observed drug concentrations and the calculated according to preset objective function is evaluated quantitatively. The iteration continues until the objective function reaches the preset values. According to Eq. (16.1), the response, $C(t)$, can be obtained given the UIR, $C_\delta(t)$, and the input, $f(t)$. Since the UIR defines the characteristic of the system, a general assumption is that the UIR would be identical for different formulations of the same compound. Therefore, explicit UIR is not required to calculate the response for one formulation $[C(t)_1]$ when another formulation has input $[f(t)_2]$ and response $[C(t)_2]$ data available.

The general numerical algorithms for convolution and deconvolution are shown as:

Convolution $C(t) = f(t) * C_\delta(t)$

$$C(t)_1 = f(t)_1 * C_\delta(t)_1 * T$$
$$C(t)_2 = [f(t)_1 * C_\delta(t)_2 + f(t)_2 * C_\delta(t)_1] * T$$
$$C(t)_3 = [f(t)_1 * C_\delta(t)_3 + f(t)_2 * C_\delta(t)_2 + f(t)_3 * C_\delta(t)_1] * T$$
$$C(t)_n = [f(t)_1 * C_\delta(t)_n + f(t)_2 * C_\delta(t)_{n-1} + \cdots + f(t)_n * C_\delta(t)_1] * T$$

Deconvolution $f(t) = C(t) // C_\delta(t)$

$$f(t)_1 = (C(t)_1/T)/C_\delta(t)_1$$
$$f(t)_2 = (C(t)_2/T - f(t)_1 * C_\delta(t)_2)/C_\delta(t)_1$$
$$f(t)_3 = (C(t)_3/T - f(t)_1 * C_\delta(t)_3 - f(t)_2 * C_\delta(t)_2)/C_\delta(t)_1$$
$$f(t)_n = (C(t)_n/T - f(t)_1 * C_\delta(t)_n - f(t)_2 * C_\delta(t)_{n-1} - \cdots - f(t)_{n-1} * C_\delta(t)_2)/C_\delta(t)_1$$

Deconvolution $C_\delta(t) = C(t) // f(t)$

$$C_\delta(t)_1 = (C(t)_1/T)/f(t)_1$$
$$C_\delta(t)_2 = (C(t)_2/T - f(t)_2 * C_\delta(t)_1)/f(t)_1$$
$$C_\delta(t)_3 = (C(t)_3/T - f(t)_2 * C_\delta(t)_2 - f(t)_3 * C_\delta(t)_1)/f(t)_1$$
$$C_\delta(t)_n = (C(t)_n/T - f(t)_2 * C_\delta(t)_{n-1} - f(t)_3 * C_\delta(t)_{n-2} - \cdots - f(t)_n * C_\delta(t)_1)/f(t)_1$$

The algorithms use piecewise integration to decompose Eq. (16.1). For example, the top block (convolution) uses pieces of the input and UIR functions, which are in reversed positions in the interval $0 \ldots t$ by assuming that both functions, input and UIR, are known as staircase functions, and both are given with a regular and common time interval, T, which requires all raw data points to be consistent. If this condition is not met, interpolations/extrapolations are needed to make the data consistent with T. The top block algorithm is most convenient for numerical calculations, especially when the inverse equations are employed in deconvolution, as shown in the midblock and the bottom block. Only for extremely small time intervals, all function values may be used just as points of the relevant time function. Otherwise it is essential to interpret them consistently as either "point" or "areas" representative for each time interval. Langenbucher presented an example dataset, which was used by several authors[44,46,54] to illustrate the convolution/deconvolution algorithms.[39] It is reproduced in Table 16.2 for reference. The data include the UIR, input, and response where T is 0.5 or 1.0.

Based on the numerical algorithm, the data are used to illustrate the convolution and deconvolution calculations as shown in Table 16.3. The top block is the convolution calculation. The next two blocks are deconvolutions for calculating the inputs and UIR, respectively.

Although the numerical calculations are rather tedious, they are usually done by computer programs in actual applications. It is helpful to understand the underlying operations, especially when a customized program is needed. The commercial software commonly used to perform convolution and deconvolution calculations include IVIVC Toolkit for Phoenix, WinNonlin by Pharsight Corporation, and IVIVCPlus, an add-on module of GastroPlus by Simulations Plus, Inc. MS Excel has also been shown to be a useful tool for IVIVC applications.[55]

16.2.2.1.3 Model-dependent deconvolution

Two commonly used deconvolution methods for estimating the apparent in vivo drug absorption profiles following oral administration of a dosage form are Wagner–Nelson and Loo–Riegelman methods.[56] These are model-dependent approaches based on mass balance. The Wagner–Nelson equation is derived from

TABLE 16.2 An Example for the UIR, Input, and Response

Time (h)	UIR			Input				Response
	$C_\delta(t)$	$C_\delta(t)_{0.5}$	$C_\delta(t)_{1.0}$	$f(t)$	$f(t)_{0.5}$	$f(t)_{1.0}$	$\int f(t)dt$	$C(t)$
0	20.227	–	–	0.41	–	–	0	0
0.5	16.364	18.296		0.334	0.372		0.185	3.36
1	13.549	14.957	16.888	0.272	0.303	0.341	0.336	5.487
1.5	11.481	12.515		0.221	0.247		0.459	6.774
2	9.946	10.714	11.748	0.18	0.202	0.227	0.56	7.492
2.5	8.792	9.368		0.147	0.164		0.641	7.831
3	7.91	8.351	8.928	0.119	0.134	0.151	0.708	7.921
3.5	7.225	7.568		0.097	0.108		0.762	7.851
4	6.681	6.953	7.296	0.079	0.088	0.099	0.806	7.681
4.5	6.239	6.46		0.064	0.072		0.842	7.452
5	5.873	6.056	6.277	0.052	0.058	0.066	0.871	7.191
5.5	5.561	5.717		0.042	0.047		0.895	6.916
6	5.29	5.426	5.582	0.035	0.039	0.044	0.915	6.639

Reproduced from Cutler DJ. Numerical deconvolution by least squares: use of polynomials to represent the input function. J Pharmacokin Biopharm 1978;6:243–263.

TABLE 16.3 Illustration of the Numerical Algorithm for Convolution and Deconvolution Based on the Data in Table 16.2

Δt	Time	Calculation and calculated results		True
1	1	$C(t)_1 = 0.341 \cdot 16.888 =$	5.759	5.487
	2	$C(t)_2 = 0.341 \cdot 11.748 + 0.227 \cdot 16.888 =$	7.840	7.492
0.5	0.5	$C(t)_1 = (0.372 \cdot 18.296) \cdot 0.5 =$	3.403	3.36
	1	$C(t)_2 = (0.372 \cdot 14.957 + 0.303 \cdot 18.296) \cdot 0.5 =$	5.554	5.478
	1.5	$C(t)_3 = (0.372 \cdot 12.515 + 0.303 \cdot 14.957 + 0.247 \cdot 18.296) \cdot 0.5 =$	6.853	6.774
	2	$C(t)_4 = (0.372 \cdot 10.714 + 0.303 \cdot 12.5\ \mathrm{I}\ 5 + 0.247 \cdot 14.957 + 0.202 \cdot 18.296) \cdot 0.5 =$	7.584	7.492
1	1	$f(t)_1 = 5.487/16.888 =$	0.325	0.341
	2	$f(t)_2 = (7.492 - 0.325 \cdot 11.748)/16.888 =$	0.218	0.227
0.5	0.5	$f(t)_1 = 3.360/(18.296\ 0.5) =$	0.367	0.372
	1	$f(t)_2 = (5.487 - 0.367 \cdot 14.957 \cdot 0.5)/(18.296 \cdot 0.5) =$	0.300	0.303
	1.5	$f(t)_3 = (6.774 - 0.367 \cdot 12.515 \cdot 0.5 - 0.300 \cdot 14.957 \cdot 0.5)/(18.296 \cdot 0.5) =$	0.244	0.247
	2	$f(t)_4 = (7.492 - 0.367 \cdot 10.714 \cdot 0.5 - 0.300 \cdot 12.515 \cdot 0.5 - 0.244 \cdot 14.957 \cdot 0.5)/(18.296 \cdot 0.5) =$	0.199	0.202
1	1	$C_\delta(t)_1 = 5.487/0.341 =$	16.091	16.888
	2	$C_\delta(t)_2 = (7.492 - 16.091 \cdot 0.227)/0.341 =$	11.259	1.1748
0.5	0.5	$C_\delta(t)_1 = 3.360/(0.372 \cdot 0.5) =$	18.065	18.296
	1	$C_\delta(t)_2 = (5.487 - 18.065 \cdot 0.303 \cdot 0.5)/(0.372 \cdot 0.5) =$	14.786	14.957
	1.5	$C_\delta(t)_3 = (6.774 - 18.065 \cdot 0.247 \cdot 0.5 - 14.786 \cdot 0.303 \cdot 0.5)/(0.372 \cdot 0.5) =$	12.381	12.515
	2	$C_\delta(t)_4 = (7.492 - 18.065 \cdot 0.202 \cdot 0.5 - 14.786 \cdot 0.247 \cdot 0.5 - 12.381 \cdot 0.303 \cdot 0.5)/(0.372 \cdot 0.5) =$	10.568	10.714

Reproduced from Cutler DJ. Numerical deconvolution by least squares: use of polynomials to represent the input function. J Pharmacokin Biopharm 1978;6:243–263.

a one-compartment model and the mass balance, $X_a = X_t + X_e$, where X_a, X_t, and X_e are amounts of drug absorbed, in the body and eliminated at time t, respectively. By derivation, the amount of drug absorbed up to time T, $(X_a)_T$, is given by: $(Xa)_T = VC_T + kV\int_0^T Ctdt$, where V is the volume of central compartment, C_T is concentration of drug in the central compartment at time T, and k is the first-order elimination rate constant. In the study of IVIVC, this is often expressed in terms of fraction (F) of the dose (D) absorbed for comparison with fraction released in vitro:

$$F_a(T) = \frac{(Xa)_T}{(Xa)_\infty} = \frac{C + k\int_0^T Ctdt}{k\int_0^\infty Ctdt} \tag{16.7}$$

where $F_a(T)$ or FD is the fraction of the bioavailable drug absorbed at time T. It should be noted that Eq. (16.7) is identical to Eq. (16.6). Therefore, the Wagner–Nelson method represents a special case of deconvolution with a single-exponential disposition. When intravenous data are not available, the apparent in vivo fractional absorption profile can be obtained by using terminal phase elimination rate constant, k, and partial areas under the plasma concentration curve using Eq. (16.7). However, it should be pointed out that: (1) k value should be derived from the true elimination phase, which may be difficult for drugs with a prolonged absorption phase and/or long half-life; and (2) only apparent absorption is estimated using this method.

The approximate equation used in absorption analysis for the two-compartment model was first published by Loo and Riegelman in 1968.[57] Wagner published an Exact Loo–Riegelman method for a multicompartment model in 1983.[58] It is a general equation for the absorption analysis of one- to three-compartment models. It requires IV data for the calculation of absorption profiles. For biexponential disposition, mass balance leads to: $(X_a)_T = X_c + X_p + X_e$, where X_c and X_p are amounts of a drug in the central and peripheral compartments at time T, respectively. By derivation,[58] $(X_a)_T$ can be determined:

$$\frac{(Xa)_T}{Vc} = C_T + k_{12}e^{-k_{21}T}\int_0^T C_t e^{-k_{21}t}dt + k_{10}\int_0^T C_t dt \tag{16.8}$$

where k_{12}, k_{21}, and k_{10} are the microconstants that define the rates of transport between compartments. On the basis of mass balance, $(X_a)_T = X_c + Xp_1 + Xp_2 + X_e$, a similar equation can be derived for triexponential disposition. The corresponding Exact Loo–Riegelman equations are given as:

$$\frac{(Xa)_T}{Vc} = C_T + k_{12}e^{-k_{21}T}\int_0^T C_t e^{-k_{21}t}dt + k_{13}e^{-k_{31}T}\int_0^T C_t e^{-k_{31}t}dt + k_{10}\int_0^T C_t dt \tag{16.9}$$

It can be shown that the Loo–Riegelman method is also a special case of deconvolution where in vivo disposition is described by two or three exponentials.[50] The theoretical and practical aspects of absorption analysis using model-dependent approaches have been thoroughly discussed by Wagner.[56]

16.2.2.2 *Mean time parameters used in Level B correlation*

Level B correlation is based on correlating mean time parameters that characterize the in vitro and in vivo time courses, for example, the in vitro or in vivo mean dissolution time (MDT) and in vivo MRT. Mean time parameters have been commonly utilized in pharmacokinetic studies and used to describe in vitro release. They are useful in studying specific models as well as less differentiated, more general system models. Many important concepts, definitions, and computations on this subject have been thoroughly discussed by Veng-Pedersen[59] and Podczeck.[60]

16.2.2.2.1 In vivo parameters

By definition, MRT is the average total time a drug molecule spends in the introduced kinetic space. It depends on the site of input and the site of elimination. When the elimination of the molecule follows first-order kinetics, its MRT can be expressed by[59]:

$$\mathrm{MRT} = \frac{\int_0^\infty tCtdt}{\int_0^\infty Ctdt} = \frac{\mathrm{AUMC}}{\mathrm{AUC}} \tag{16.10}$$

where AUMC is area under the moment curve. Estimates for MRT can be calculated by fitting $C(t)$ to a polyexponential equation followed by integration or by using trapezoidal rules.

For noninstantaneous input into a kinetic space, such as oral absorption, the MRT estimated from extravascular data includes a contribution of the mean transit time for input, known as mean absorption time (MAT, or mean arrival time, or mean input time).[59]

The MAT of drug molecules represents the average time taken to arrive in that space, and it can be estimated as:

$$\mathrm{MAT} = \frac{\int_0^\infty t f_{\mathrm{in}}(t)dt}{\int_0^\infty f_{\mathrm{in}}(t)dt} = \frac{\mathrm{AUMC}}{\mathrm{AUC}} \tag{16.11}$$

where $f_{\mathrm{in}}(t)$ denotes an arbitrary rate of input into the kinetic space. For oral delivery, the MAT can be determined according to the equation:

$$\mathrm{MAT} = \mathrm{MRTpo} - \mathrm{MRTiv} \tag{16.12}$$

The term MAT thus obtained represents the mean transit time involved in apparent absorption process in the gastrointestinal (GI) tract. When the formulation contains a solid drug, the MAT includes in vivo dissolution as well as absorption. If data of the same drug given in a solution state are available, the in vivo MDT can be estimated by:

$$\mathrm{MDT}_{\mathrm{solid}} = \mathrm{MAT}_{\mathrm{solid}} - \mathrm{MAT}_{\mathrm{soln}} = \mathrm{MRT}_{\mathrm{solid}} - \mathrm{MRT}_{\mathrm{soln}} \tag{16.13}$$

16.2.2.2.2 In vitro parameters

The measured amount of a drug substance in a cumulative-release profile can be considered as a probability that describes the time of residence of the drug substance in the dosage form. Therefore, a dissolution profile may be regarded as the distribution function of the residence times of each drug molecule in the formulation.[49] By definition, the MDT is the arithmetic mean value of any dissolution profile. If the amount of the drug remaining in the formulation is plotted as a function of time, the arithmetic mean value of the residence profile is the MRT of the drug molecules in the dosage form.

The techniques that are used to calculate MDT or MRT can be divided into model-independent (pragmatic plane geometry and prospective area) and model-dependent methods (eg, polyexponential, Weibull, and overlapping parabolic integration).[49] In general, model-independent approaches are used when release kinetics are unknown. These methods are based on area calculations from the amount released at various times. The following simple method is often used to determine the MDT and MRT using trapezoidal rules[49]:

$$\mathrm{MDT} = \frac{\int_0^\infty (M_{\max} - M(t))dt}{M_{\max}} = \frac{\mathrm{ABC}}{M_{\max}} \tag{16.14}$$

$$\mathrm{MRT} = \frac{\int_0^\infty tA(t)dt}{\int_0^\infty A(t)dt} \tag{16.15}$$

where ABC is the area between the drug dissolution curve and its asymptote. $A(t)$ is the amount of drug remaining in the dosage form at time t. $M(t)$ and $M_{\max}$ are the amount of drug released at time t and the maximal amount released, respectively. The model-dependent methods are based on the derived parameters of functions that describe the release profiles. It should be noted that one important source of errors in calculations comes from the often incomplete release. The calculation of the moments in such case is based on the maximum drug release. For systems that have a complete drug release, the size of errors depends on the number of data points and the curve shape.[49]

16.2.2.3 Summary parameters used in Level C correlation

The extent and rate of drug release from a dosage form are often characterized by one or more of the single measurements (eg, Q_{60}, $T_{50\%}$, or $T_{85\%}$), particularly when there are not enough data points available to define the time functions of the profiles, or there are simply no suitable models that describe the dissolution curves. These parameters are most often obtained either directly from the dissolution measurements or by interpolation. Although they do not adequately characterize the whole dissolution process, they are utilized in quality control and in Level C correlation studies. The in vivo parameters used to correlate with the in vitro parameters are bioavailability parameters reflecting the rate and extent of absorption (eg, AUC and $C_{\max}$).

16.2.2.4 Establishment of a Level A IVIVC model

Establishing an IVIVC model requires in vitro data from formulations with different release rates (eg, fast, medium, and slow) and a discriminating in vitro test methodology. The corresponding in vivo response can be plasma concentrations or the amount of drug released and/or absorbed in vivo. The latter is obtained from the observed plasma concentration-time curve by deconvolution. There are advantages and disadvantages for either type of response variable. When plasma concentration is used as a response variable (single-stage approach), the link between the in vitro dissolution profile with the in vivo plasma concentration profile has clear clinical relevance because many pharmacokinetic parameters such as $C_{\max}$, $T_{\max}$, and

AUC are directly derived from the plasma concentration-time profile. Using the amount of drug released/absorbed as a response variable (two-stage approach) is intuitively straightforward because the in vitro and in vivo parameters are directly compared.

16.2.2.4.1 Two-stage approach

A deconvolution-based IVIVC model is established using a two-stage approach that involves an estimation of the in vivo release/absorption profile from the plasma concentration-time data using an appropriate deconvolution technique (eg, Wagner–Nelson, numerical deconvolution) for each formulation. Subsequently, the calculated in vivo percent absorbed or released is correlated with the percent released in vitro, as illustrated in Fig. 16.3 using a basic linear model with intercept (a) and slope (b):

$$(\%\ \text{absorbed})_{\text{in vivo}} = a + b(\%\ \text{released})_{\text{in vitro}} \qquad (16.16)$$

A slope closer to one indicates a 1:1 correlation, and a negative intercept implies that the in vivo process lags behind the in vitro dissolution. A positive intercept has no clear physiological meaning. It can be a result of relatively high variability or curvature at the early time points. When the in vitro data are not in the same time scale as the in vivo absorption, it is usually necessary to incorporate a scaling factor, such as time-shifting and time-scaling parameters, within the model. Nonlinear models, while uncommon, may also be appropriate.[13] The two-stage approach is the most frequently used in building IVIVC models.

For an IVIVC with a two-stage approach, an important aspect is to find the relationship between the in vitro dissolution and in vivo dissolution/absorption profiles. The general strategy is to use scaling and/or shifting or have the time scaled and shifted to match the in vitro and in vivo profiles. In the software commonly used, such as Winnonlin Phoenix, the following models are built in.

A. Fabs = Diss(Tvivo)
B. Fabs = AbsScale * Diss(Tscale * Tvivo)
C. Fabs = AbsScale * Diss(Tscale * Tvivo − Tshift)
D. Fabs = AbsScale * (Diss(Tscale * Tvivo − Tshift) − AbsBase)

Where Fabs is the in vivo absorption fraction; Diss() is a function for dissolution, dependent on the in vitro time in the parenthesis, which could be the same as in vivo time: Tvivo as in model A, the scaled in vivo time by Tscale as in B, or the in vivo time scaled by Tscale and shifted by Tshift as in C and D. In B, C, and D, to get the in vivo fraction (Fabs), the in vitro dissolution (Diss) has to be scaled by AbsScale. Before this scaling, in D, Diss would be corrected by baseline (AbsBase).

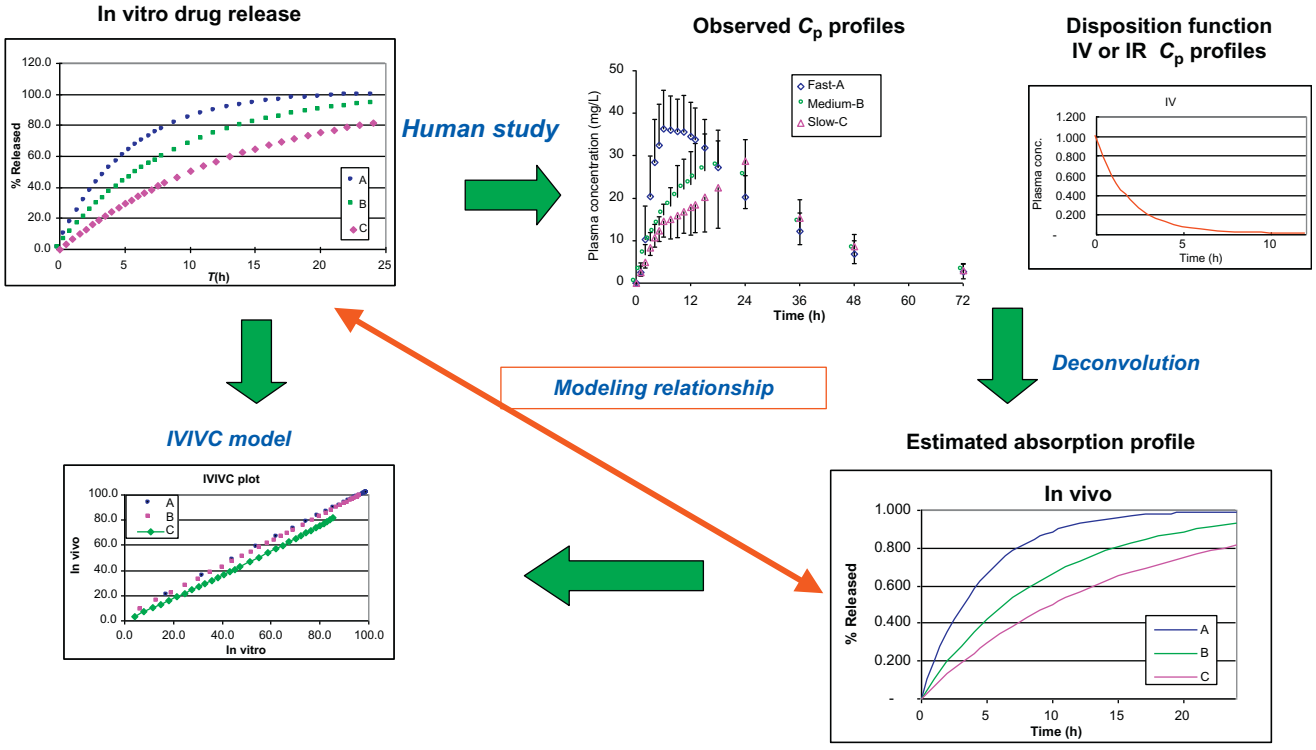


FIGURE 16.3 Illustration: building a Level A IVIVC model using the two-stage approach.

When all the scaling and shifting factors are included, it may be difficult to identify the appropriate model. To get an insight of the effect of these factors, a simulation study was performed in which the effects of three levels of each of the scaling (scale, AbsScale in B, C and D), baseline correction (shift, AbsBase in D), time scaling (Tscale), and time shifting (Tshift) were examined. The levels for the four parameters are provided below:

Level	Scale	Shift	Tscale	Tshift
1	0.3	−30	0.5	−5
2	1	0	1	0
3	1.5	30	2	5

There were a total of 81 possible combinations, and the results are shown in Fig. 16.4, where the values of the four factors are labeled in each panel.

Panel 41 shows case A, where the in vitro time is the same as in vivo time, resulting in a 1:1 relationship. The in vitro and in vivo profiles are superimposed. When the model is baseline corrected, the shape of the profiles does not change, and the profile just shifts up (or down), while scaling gets an n:1 relationship (n is dependent on the values of AbsScale). The time scaling and shifting lead to the inconsistent in vitro time and in vivo time and different shapes of the profiles. Time scaling and shifting are needed if the in vitro and in vivo profiles exhibit different rates and/or a difference in lag time.

In addition to these built-in models, the software usually allows the user to specify a customized model. As shown in Fig. 16.5, the relationship between the in vivo and in vitro profiles seems to follow a sigmoidal shape, and an E_{max} model could be used.

In summary, the patterns of in vitro and in vivo profiles may provide clues for an appropriate model. Plotting the in vitro and in vivo data at the same time points is helpful for recognition of the patterns in addition to the Levy plots.

16.2.2.4.2 Single-stage approach

An alternative modeling approach based on convolution can be utilized to directly predict the time course of plasma concentrations using Eq. (16.1) in a single step.[61] Based on the assumption of equal or similar release rates between in vitro and in vivo, the input rate, $f(t)$, is modeled as a function of the in vitro release data with or without time scaling to predict the in vivo plasma profiles by convolution with the dose-normalized plasma data from an IV or IR reference dose. The IVIVC is assessed and validated by statistically comparing the predicted with the observed plasma levels. This convolution-based modeling focuses on the ability to predict measured quantities rather than indirectly estimated in vivo fraction absorbed and/or released. Thus, the results are more readily evaluated in terms of the effect of in vitro release on in vivo performances, for example, AUC, $C_{max,}$ and duration above minimum effective concentrations. For instance, in using this approach to estimate the plasma concentrations from the in vitro data, a poly-exponential UIR with lag time could be used in the model as follows:

$$C_\delta(t) = \sum_{i=1}^{nex} A_i{}^* e^{-\alpha_{(i)}(t-t_{lag})} \tag{16.17}$$

where, nex is the number of exponential terms in the model, t_{lag} is the absorption lag time, and C_t is the plasma concentration at time t. The input rate may be modeled as a function of the in vitro cumulative amount dissolved. For example, Veng-Pedersen, et al. reported a scaled convolution-based IVIVC approach by which the dissolution rate curve was first obtained via differentiation of a monotonic quadratic spline fitted to the dissolution data. Using time and magnitude scaling, the dissolution curve was then mapped into a drug concentration curve via a convolution by a single exponential and the estimated UIR function. The model was tested by cross-validation and demonstrated to be predictive of the systemic drug concentration profiles from the in vitro release dissolution data using four different tablet formulations of carbamazepine.[62]

It should be noted that the single-stage approach is based on the assumption of an LTI relationship between the input (drug release) and the response (plasma concentrations). Multiple formulations with different release rates are usually used in establishing an IVIVC. If a significant fraction of the dose of the slow-releasing formulations is released beyond the site (s) of drug absorption (ie, truncated absorption), an overestimation of plasma concentrations can occur because Eq. 16.1 predicts the same dose-normalized AUC as the reference dose used to estimate $C_\delta(t)$.[62] To address the potential discrepancies between in vitro and in vivo release/absorption, Gillespie proposed an extended convolution-based IVIVC model using a function relating the cumulative amounts released or the release rate in vitro (x_{vitro}) to that in vivo (x_{vivo}), $x_{vivo} = f\ (x_{vitro})$. Thus, plasma concentrations of multiple ER formulations can be more accurately predicted by substituting $f(t)$ with x_{vivo} in Eq. (16.1) if there is an IVIVC. Selection of a specific functional form of x_{vivo} can be based on a mechanistically understanding of the in vitro/in vivo relationship or semiempirically based on the goodness of model fitting. Certain plausible relationships include linear, nonlinear, or

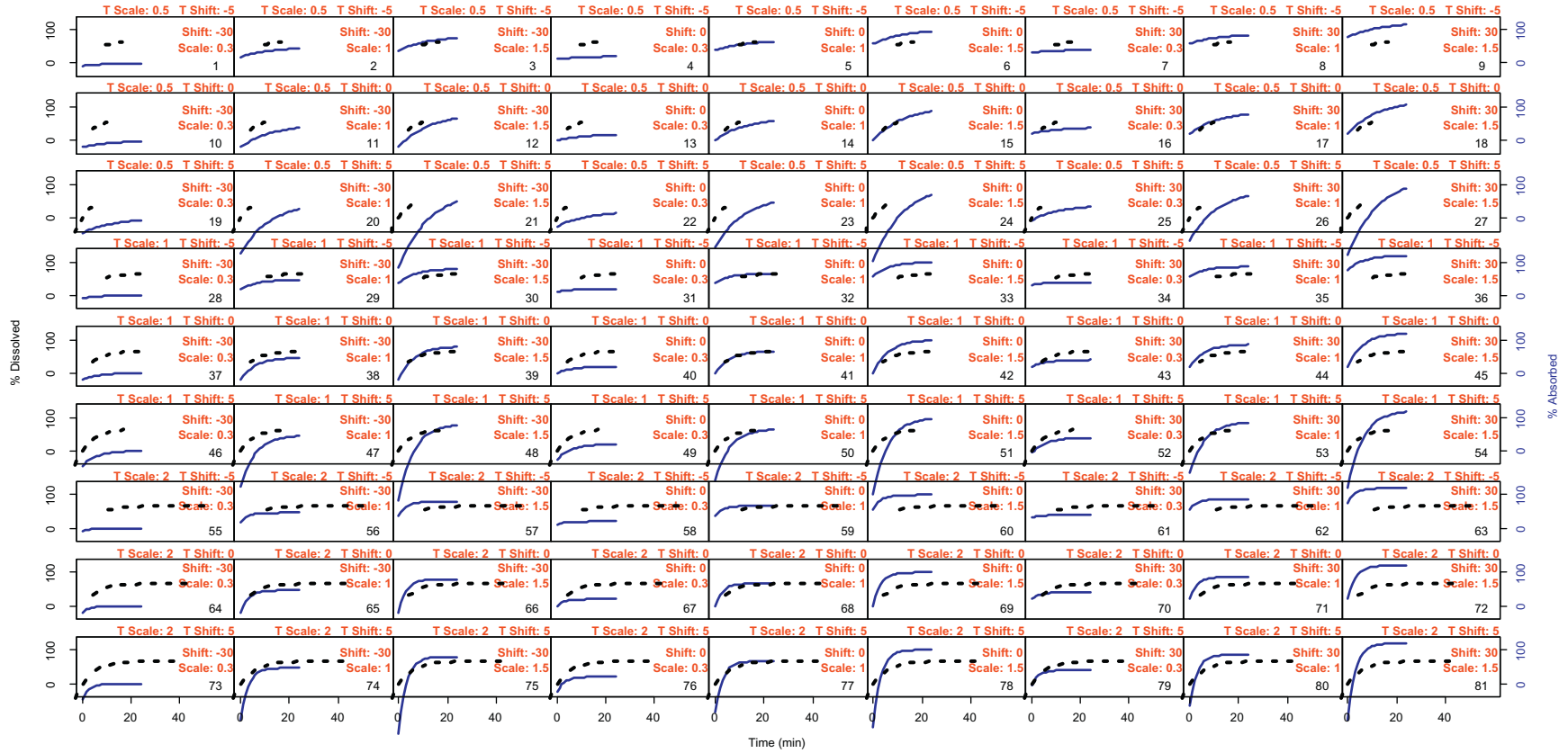


FIGURE 16.4 Simulated in vitro and in vivo profiles based on Model D Fabs = AbsScale * (Diss(Tscale *Tvivo − Tshift) − AbsBase) (dashed and solid lines are the in vitro dissolution and in vivo absorption profiles).

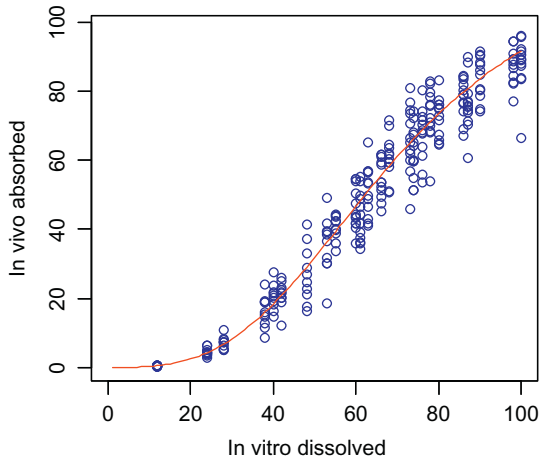


FIGURE 16.5 A nonlinear in vitro/in vivo relationship fit by E_{max} model.

time-variant functions; linear or nonlinear time-scaling for taking into account the effects of lag time; truncated absorption; or saturable presystemic metabolism.[62,63] For ER systems, the apparent absorption in vivo is limited by the drug release from the dosage form, of which the kinetics is determined by the system design. In many cases, the release kinetics can be described by one of the following models: zero-order, first-order, square root of time, or Peppas's exponent models.[64] Therefore, the parametric function, $C(t)$, describing the plasma concentrations of an ER dosage form can be defined via convolution of $C_\delta(t)$ with the input, $f(t)$, prescribed from the in vitro model according to Eq. (16.1). For instance, the functional form of $C(t)$ can first be obtained by convolution of a prescribed function of input (eg, first-order) with a known UIR. The unknown parameters remaining in the $C(t)$ equation are those from the prescribed input function, $f(t)$, which can then be solved by fitting the $C(t)$ equation to the observed plasma concentrations. The resulting $f(t)$ is compared with the observed in vitro data to evaluate IVIVC. This approach is also known as DTC, as discussed in the previous section.

In predicting $C(t)$ by convolution, data from an IV or oral solution is desirable because it provides an estimate of $C_\delta(t)$ independent of the ER data. However, such a reference dose is not always available, particularly for compounds having low aqueous solubility. Nevertheless, an estimation of $C(t)$ by convolution for evaluation of IVIVC is still possible using only data from ER formulations.[62] In such cases, the prescribed parametric functional forms of both $C_\delta(t)$ and $f(t)$ can be mechanistically or empirically selected and substituted into Eq. (16.1). The parameters of $C_\delta(t)$ are then estimated by fitting the overall convolution model to the plasma concentrations of ER formulations. Predictive performance of the IVIVC is evaluated by comparing the predicted and observed results. It should be pointed out that the ability of the model to predict changes of the in vivo plasma concentrations with varying release rates should be validated by separately or simultaneously fitting the data from multiple formulations. By doing so, a $C_\delta(t)$ function can be reliably defined. Thus, one of the most critical requirements of this approach is to use at least two ER formulations with different release rates in the assessment of IVIVC.[61]

While the use of the two-stage procedure is more widespread, the convolution approach has gained increased interest. O'Hara et al. compared odds and identity models that include a convolution step using the data sets of two different products and a nonlinear mixed-effects model fitting software to circumvent the unstable deconvolution problem of the two-stage approach.[65] Gaynor et al. used a simulation study to show the convolution modeling approach produces more accurate results in predicting the observed plasma concentration-time profile.[66] Jacobs et al. described an IVIVC model for an oral product consisting of IR and ER components of galantamine by combining the IR and ER pharmacokinetic profiles using a one-stage convolution-based method.[67] The average percentage prediction error (PE) indicated a good fit of the new model.

16.2.2.4.3 Compartmental and population approach

Both convolution- and deconvolution-based methods assume that the system being modeled is linear, but in practice, this is not always the case, as many drugs are eliminated by nonlinear (saturable) processes.[68–70] Therefore, a linear system approach can fail when the disposition of a drug substance falls within a nonlinear range. The extent of this failure depends on the magnitude of the nonlinearity observed.[71] To overcome problems resulting from nonlinear kinetics, the use of a method that can reliably accommodate the nonlinear characteristic is imperative. One of the effective methods to analyze nonlinear data is to use a compartmental approach. In a published study, the performance of a compartmental approach for nonlinear systems using the population approach for IVIVC modeling is evaluated.[72]

The population approach is to seek the sources and correlation of the variability in vivo, in vitro, and their relationship. The principles are similar to population pharmacokinetics (popPK), and understanding of the basic popPK concept will be greatly helpful. Fig. 16.6 introduces the concept of two types of variability considered by popPK. The solid circles are the data points for subject 1, from which a fit line is generated (dashed line). The open circles are the data points for all other

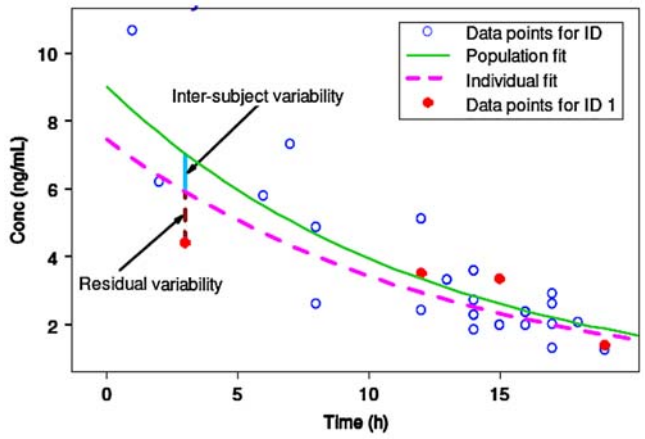


FIGURE 16.6 An illustration of popPK concept.

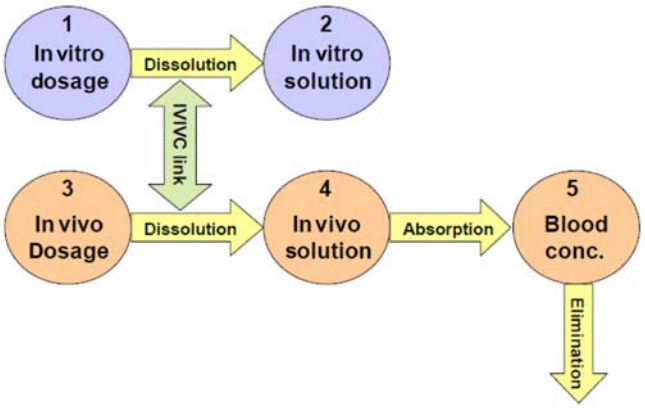


FIGURE 16.8 IVIVC based on a compartmental approach.

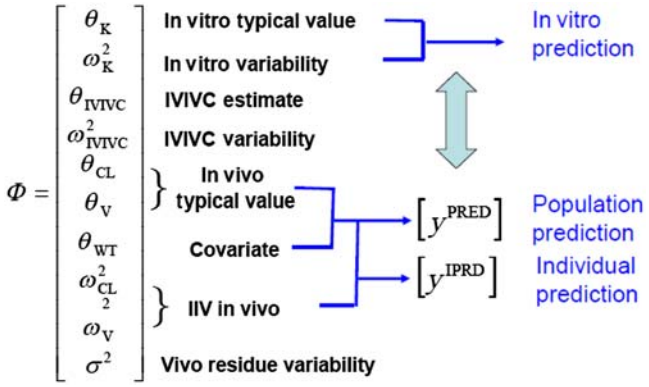


FIGURE 16.7 Various sources of variability considered in population IVIVC.

subjects. A population fit line (solid line) is generated from all the points (both solid and open circles). If one data point at a specific time (3.5 hours) for a specific subject (ID 1) is considered, conceptually, the PE is composed of two parts: one resulting from interindividual variability and another from residual variability.

For population IVIVC, not only the residue variability and the intersubject variability but also the in vitro variability and the variability of in vitro/in vivo correlation should be considered, as shown in Fig. 16.7. Clearance (CL) and volume distribution (V) are different from their typical values (θ_{CL} and θ_V) due to the fixed effects such as weight (θ_{WT}), age (θ_{AGE}), and gender (θ_{SEX}). They also can be affected by the intersubject variability (ω^2_{CL} and ω^2_V) and the residue variability (σ^2). In addition, the fixed effect (θ_K) and the random effect (ω^2_K) for in vitro (dissolution) performance and the fixed effect (θ_{IVIVC}) and the random effect (ω^2_{IVIVC}) for the correlation should be taken into consideration.

Fig. 16.8 shows how an IVIVC relating to in vitro and in vivo drug dissolution could be described in terms of compartments.[73] This technique retains most of the benefits of the convolution approach: in vitro fractions dissolved and in vivo plasma concentrations are modeled directly in a one-step process, and the IVIVC can be specified to incorporate random effects, time dependence, and scale factors as required by each particular set of data.

16.2.2.5 *Establishment of a Level C IVIVC model*

Building a Level C IVIVC model is rather straightforward. It involves correlating the amount dissolved at various time points with C_{max}, AUC, or other suitable bioavailability parameters. Data from at least three formulations of different dissolution rates are required for establishing a linear or nonlinear relationship between the in vitro and in vivo parameters because each data point of the correlation plot corresponds to only one formulation. A single-point Level C correlation may facilitate formulation development or allow a dissolution specification to be set at the specified time point. The information is generally insufficient for justifying a waiver of a bioequivalence study. A multiple-point Level C correlation may be useful to support a biowaiver if the correlation has been established over the entire dissolution profile with one or more bioavailability parameters of interest. A relationship should be demonstrated at each time point with the same parameter such that the effect on the in vivo performance of any change in dissolution can be assessed. When such a multiple Level C correlation is achievable, the development of a Level A correlation is often more likely. A multiple Level C correlation should be based on at least three dissolution time points covering the early, middle, and late stages of the dissolution profile.[2]

16.2.3 Evaluation of a correlation

To assure a useful and reliable Level A IVIVC, the FDA guidance requires that the model should be demonstrated consistently with two or more formulations with different release rates. Therefore, the first important step in evaluating the appropriateness of the model is to test whether a single correlation fits all tested formulations. One of the statistical assessment approaches is to compare the fit of a reduced model where all tested formulations are fitted to a single correlation line with that of a full model, where each formulation is fitted to a different correlation line.[33] If both models fit well, the IVIVC is considered validated. If the full model fits well, but the reduced model does not, or if the full model is statistically different from the reduced model at significance level of 0.05, the IVIVC becomes formulation dependent and, therefore, is invalid. It is noted that the time-scaling factor, if used, should also be the same for all formulations. Different time scales for each formulation indicate the absence of an IVIVC.[2]

Following the establishment of a Level A IVIVC model, it is necessary to demonstrate that prediction of the in vivo performance from the in vitro dissolution characteristics is accurate and consistent. The FDA guidance suggests evaluating the goodness of fit by measuring the prediction error (PE), that is, differences between the observed and the predicted values over a range of in vitro release rates. As illustrated in Fig. 16.9, determination of PE involves calculation of the in vivo absorption (input) profiles from the in vitro data using the established IVIVC model, followed by the prediction of the corresponding plasma concentration profiles via convolution. The guidance further elaborates approaches to validate the model internally and externally. Internal validation can be accomplished through measuring PE using data from the same study used to develop the IVIVC. The internal PE evaluates how well the model describes the data used to develop the IVIVC. It could be adopted for cases where the IVIVC was derived using two or more formulations with different release rates, providing the drug is not considered a narrow therapeutic index drug. The external validation approach requires a data set that was not used in the development of the IVIVC, such as formulations with a different release rate, formulations with minor manufacturing process changes, or a formulation from a different manufacturing batch obtained from a different study. It is desirable and affords greater confidence in the model.

The criteria used in the FDA guidance on IVIVC are as follows: For predicted C_{max} and AUC, the mean absolute percent prediction error (% PE) should not exceed 10%, and the PE for individual formulations should not exceed 15%. A PE of 10–20% indicates inconclusive predictability and illustrates the need for further study using additional data sets. For drugs

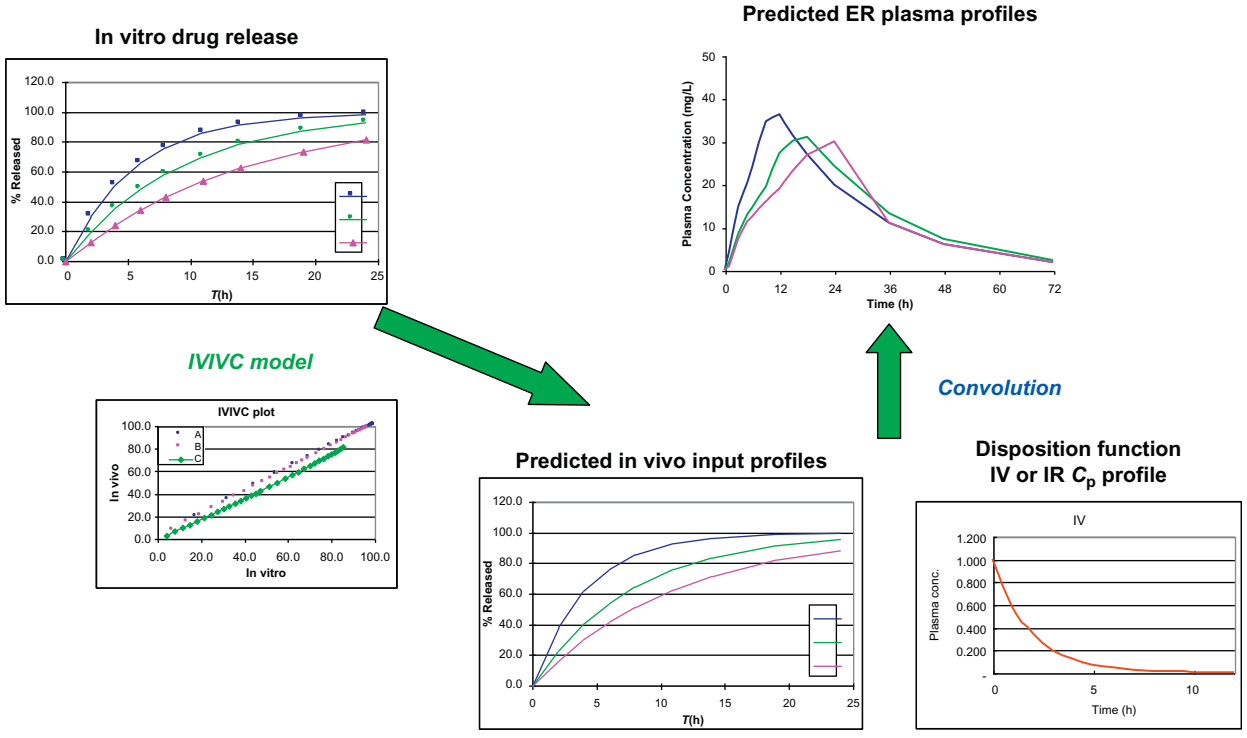


FIGURE 16.9 Illustration: prediction in IVIVC model validations and applications.

with a narrow therapeutic index, external validation is required despite acceptable internal validation, whereas internal validation is usually sufficient with nonnarrow therapeutic index drugs. In general, the less data available for the initial IVIVC development and evaluation of predictability, the more additional data may be needed to define completely the IVIVC's predictability. A combination of three or more formulations with different release rates is considered optimal.[2]

For Level C correlations, assessment of the predictability will depend on the type of application for which the correlation is to be used. The methods and criteria are the same as those for a Level A correlation. A recent example of the successful development and validation of a Level C IVIVC was reported by Kesisoglou et al. for suvorexant IR tablets based on amorphous solid dispersion.[74] Four different batches of tablets were manufactured using a hot-melt extrusion process and compressed at different tablet hardnesses to produce distinct dissolution profiles. These batches were evaluated in a relative bioavailability study in healthy volunteers followed by investigation of a relationship between dissolution and C_{max}. A validated multiple Level C IVIVC was developed between C_{max} and Q_{10}, Q_{15}, Q_{20}, Q_{30}, Q_{45} and the disintegration time, and it was used to set clinically relevant dissolution specification and in process control of tablet hardness.

16.3 CONSIDERATIONS IN IVIVC DEVELOPMENT

Defining a quantitative and reliable relationship between in vitro drug release and in vivo absorption is highly desired for the rational development, optimization, and evaluation of solid dosage forms and the manufacturing process. A validated IVIVC can significantly (1) increase the development efficiency by reducing the time and resources required for formulation and process development, scale-up, and optimization; (2) assure product quality by setting meaningful specifications; and (3) reduce regulatory burdens by using an in vitro test as a surrogate to in vivo bioavailability studies required for certain postapproval changes. Therefore, exploring and developing IVIVC where possible should be an important part of solid oral product development.

Since the early 1990s, extensive colloquiums and research publications have mainly focused on the basic principles, development methodology, modeling, evaluation, and applications of IVIVC and have laid a solid scientific foundation for regulatory guidance and decisions. Most of the relevant scientific discussions are based on the premise of availability of in vitro and in vivo information appropriate for establishing an IVIVC. However, it should be emphasized that modeling and statistical assessment is only one aspect of IVIVC development, and obtaining suitable in vitro and in vivo data is not a given. To achieve an IVIVC, at least two formulations that differ in the in vivo and in vitro performance should be available. In fact, many failed attempts in achieving an IVIVC for solid products can usually be attributed to a lack of a predictive in vitro test or an adequate in vivo difference between test formulations. Thus, it is crucial to understand how a drug's physicochemical and biopharmaceutical properties, delivery technology, and formulation design and their interplays with both the gastrointestinal tract (GIT) and the in vitro test methodology may impact in vitro and in vivo data.

16.3.1 In vivo absorption versus in vitro test considerations

In developing IVIVC, the in vitro parameter commonly utilized is drug release, which can be determined with precision under a controlled condition. It is a function of drug and dosage form characteristics and test methods and conditions. The in vivo response is usually the dissolution and absorption estimated from the availability of the drug in systemic circulation that typically has high variability. It is a function of a multitude of physicochemical, biopharmaceutical, and physiological variables, formulation, and their interactions. For test formulations that exhibit varying apparent in vivo absorption characteristics, the most critical element in establishing an IVIVC is the ability of in vitro tests to correlate quantitatively with the in vivo performance. To assess the challenges and the opportunities of developing a predictive and meaningful in vitro test, it is essential to first understand the absorption characteristics of the drug substance and the complexity of drug absorption from dosage forms in the GI tract.

16.3.1.1 Apparent drug absorption from the GI tract

The in vivo drug release and subsequent absorption from a solid product, particularly a modified-release dosage form, take place in one of the most complex environments where the interplays of the GI tract with the dosage forms as well as with intralumenally released drugs is highly dynamic. The process is known to be influenced by (1) drug property (solubility, ionization, stability, solid phase, lipophilicity, permeability, surface area, and wetting property); (2) dosage form design (dose, release mechanism, composition, type,

size and transit characteristics of the dose unit, sensitivity to shear force, and drug release location or duration); and (3) GI physiology and biology (motility, residence time, food, lumen contents, fluid volume, transport pathways and mechanism, enterohepatic recycling, effective permeability, surface area, metabolism, uptake and efflux transporters, and microflora). This type of information is highly valuable in anticipating and deciphering in vitro and in vivo data of the drug molecule and the dosage forms as related to the development of an IVIVC. For example, an apparently truncated absorption or a significantly decreased extent or rate observed with the slow-release formulations 4–5 hours post dosing may be related to one or more the following factors discussed in chapter "Rational Design of Oral Modified-Release Drug Delivery Systems" and literature[75–81]: lower apparent permeability in the large bowel due to inadequate lipophilicity or involvement of gut enzymes/transporters; high dose/solubility ratio; microbiotic degradation or decreased delivery rate of the formulations due to a change in lumenal environment. When assessing discrepancies between in vitro and in vivo data, delivery technology and formulation design, and release mechanism, in vitro test method and condition are important factors to consider. For instance, a dosage form generally experiences a low and well-controlled flow field and shear rate in a typical in vitro test, while it is known to be subjected to a myriad of physical forces in vivo, including pressure and shear stress from cyclic strain associated with villous motility and repetitive deformation engendered by peristaltic muscular contractions and relaxation of the intestinal wall.[82] As a result, it is not uncommon to observe differences between in vitro and in vivo behaviors for formulations that are sensitive to the release environment, particularly for hydrophilic matrix formulations containing a relatively high percentage of soluble components and/or a low percentage of low-viscosity-grade polymers. For example, an appreciably more rapid in vivo release and/or increased in vivo differentiation among formulations exhibiting different in vitro release rates can be attributed to a change of the release-control mechanisms from diffusion in vitro and erosion in vivo because a predominantly diffusion-controlled drug release in vitro is known to be less sensitive to variations in gel strength and matrix integrity.[11,12,83,84] On the other hand, decreased in vivo discrimination can be a result of a weaker and more environmentally sensitive gel, active pharmaceutical ingredient (API)–polymer interaction, or nonnegligible contribution of particle dissolution to the drug release in vivo due to a nonsink condition.[85] More detailed information on variables that may influence apparent absorption, including drug properties, biopharmaceutical factors, and product/process design, can be found in chapters "Oral Absorption Basics: Pathways and Physicochemical and Biological Factors Affecting Absorption," "Oral Drug Absorption: Evaluation and Prediction," "Predictive Biopharmaceutics and Pharmacokinetics: Modeling and Simulation," "Rational Design of Oral Modified-Release Drug Delivery Systems," and "Product and Process Development of Solid Oral Dosage Forms," respectively.

16.3.1.2 In vitro test method

Among various in vitro physicochemical tests, dissolution testing is one of the most important tools for product quality assessment, process control, and for assuring sameness after making formulation or process changes. While different dissolution tests have been applied to determine drug release, including the commonly used compendia methods (eg, basket, paddle, or reciprocating cylinder), the dissolution rate of a specific dosage form is, in many cases, an arbitrary parameter that often varies with the test methodology with respect to its relevance to the in vivo performance unless an IVIVC is demonstrated.[32,86] One of the main reasons is that none of the existing in vitro methods represents or mimics the highly dynamic and complex in vivo conditions and a wide range of variables involved in the drug absorption from solid dosage forms in the GI tract.

Over the years, the exploration and the development of various in vitro tests or models have been undertaken in an attempt to match or predict in vivo data and/or to simulate specific aspects of the GI condition.[87] These studies include designs ranging from simple and very complex setups and can be divided into three broad categories listed in Table 16.4. The common attempt made in these studies is to match in vivo data by adjusting the in vitro discrimination and release rate of different test formulations through altering test media, hydrodynamics, mixing, creating shear force, or dynamic flows. It should be noted that many tests are static and only capable of imitating very limited aspects of the complex in vivo conditions. For example, some test methods are focused on test medium, supersaturation/ precipitation, and/or apparatus variables. Other models based on multiple compartments or vessels are designed to integrate dissolution with digestion and absorption or simulate motility and successive dynamic processes occurring in the GI tract. The most complicated in vitro model is a multicompartment, dynamic, computer-controlled system designed to simulate the human stomach, small intestine (TIM-1), and large intestine (TIM-2), respectively.[88,89] The system is intended for gathering information about the effects of various simulated GI conditions on biopharmaceutical behavior of the API

TABLE 16.4 Types of In Vitro Drug Release Tests for Predicting In Vivo Performance

Type	Examples	Variables altered	References
Standard pharmacopoeial methods	USP I, II, III, or IV	Hydrodynamics; agitation; pH; ionic strength; surfactants; enzymes; additives (eg, bile salts, lecithin, fat, milk)	10,93–99
Modified pharmacopoeial methods	USP II + polystyrene beads	Hydrodynamics; shear stress; mechanical attrition; mixing; pH; test medium; release vs dissolution, etc.	100–112
	USP II + Stationary basket		
	USP II + two-phase		
	Milk/fat as test medium		
	FeSSIF/FaSSIF as test medium		
	Ex vivo intestinal fluid		
Complex models	Rotating dialysis cell	pH; food; mixing; motility; transit; digestion, secretion; microflora, permeation, etc.	8,86,113–119
	Flow-through cell drop method		
	Multicompartmental systems		

and dosage forms. However, setting up experiments and generating data are time consuming and extremely labor intensive. It is unsuitable for routine product quality control. In addition, literature reports of its utility and application in IVIVC development are still absent, although there are reports of its use in evaluating relative formulation performance.[90,91]

In summary, the most important aspect of IVIVC is to develop a predictive in vitro test. In practice, product quality and performance characterized by such a test is the drug release from the dosage forms. However, the state of the art has yet to allow the development of a universal predictive model independent of drug molecule and dosage form characteristics because physiological, physiochemical, and biological conditions of the GI tract and their interplays with the API and drug products in vivo are extremely complex[7,85,92] Nevertheless, it remains possible to develop tailored predictive tests for an individual API and/or product based on the understanding of the critical aspects of importance to in vivo absorption. For instance, when in vivo drug release is the dominant controlling factor in the rate of appearance of the drug in the blood, the focus on key API and formulation variables that affect the drug release and the corresponding in vitro test method is crucial in identification of a predictive test condition. Furthermore, continued pursuit of such a test not only can bring considerable benefit to product development and quality control, but it also will help advance scientific understanding and continued effort in developing new biopharmaceutical tools for improved prediction of in vivo drug absorption.

16.3.2 Drug and formulation considerations

The essential conditions for correlating in vitro dissolution with in vivo performance include (1) the apparent in vivo absorption is dissolution-rate limited; (2) in vitro drug release (dissolution and/or erosion) is the critical dosage form attributes; (3) test formulations exhibit different in vivo performances; (4) in vitro test is discriminating (IVIVR) and/or predictive (IVIVC) of in vivo performance. When these conditions are met, an IVIVC is deemed feasible. To develop an in vitro dissolution test that can be used to evaluate how changes in formulation and/or manufacturing processes may affect in vivo performance, integrated knowledge of drug properties, delivery technology, formulation, and biopharmaceutics is required to understand the specific challenges, choose the right biopharmaceutical tools, and design the appropriate studies.

16.3.2.1 Immediate-release dosage forms

It is generally accepted that an IVIVC is more difficult to achieve for IR dosage forms. With rapid drug release, the apparent drug absorption of the dosage form usually occurs in the upper small intestine where many factors other than dissolution are known to limit, affect, or contribute to the apparent absorption, as discussed previously. The short absorption phase, in most cases, is difficult to characterize, making Level A IVIVC modeling less likely than Level C or multiple Level C. Thus, the feasibility of an IVIVC is often drug and product dependent. For BCS class 2 drugs, of which dissolution is prolonged due to very low solubility and/or product and process characteristics, an

IVIVC is possible. One such example is the relatively slow-eroding polymer matrix of amorphous solid dispersions that contain a high level of carrier polymer(s) and are manufactured using a melt-extrusion process.[74,120] In the case of BCS class 1 drugs, an IVIVC is less likely unless the drug dissolution is significantly slowed due to formulation, processing, or the compound belongs to a borderline solubility classification. IVIVC is very rare for BCS class 3 drugs, of which gastric emptying, permeability, and transporters are typically involved in drug absorption. The opportunity for IVIVR or IVIVC may sometimes exist for BCS class 4 drugs, of which dissolution, permeability, and transporters can all influence the rate of in vivo absorption depending on the relative contributions of each factor and whether the low permeability classification is borderline or due to metabolism.[121] Challenges and various physicochemical, biopharmaceutical, and physiological factors that need to be considered in developing IVIVC and IVIVR of IR oral dosage forms have been reviewed extensively.[92,122] One of the major issues with assessing the feasibility of IVIVC using the BCS classification is that the boundaries for high solubility and permeability are set very conservatively.[92] As a result, each BCS class includes compounds with a broad range of properties and potentially different rate-limiting steps for absorption. For example, a low-solubility compound that narrowly misses the 90% fabs boundary would technically fall into BCS class 4, and yet it is very unlikely to have permeability-limited absorption. For such a compound, IVIVC remains possible. Similarly, certain BCS class 2 compounds are less likely to be candidates for developing an IVIVC.[92] Therefore, it is important to understand the specific characteristics of API and drug products, including the operating principles and manufacturing process. It should be noted that when a quantitative in vitro/in vivo relationship cannot be established or validated, an IVIVR that defines a dissolution range or space with demonstrated bioequivalence can also be useful in assuring acceptable in vivo performance via mapping in vitro and in vivo data of different formulations.[84,92,123]

16.3.2.2 *Extended-release (ER) dosage forms*

Compared to IR products, an IVIVC is, in general, more desirable and readily defined for ER solid dosage forms. In fact, it is often expected that an in vitro release test is predictive or has an in vivo rank order because drug input from the GI tract is by design controlled or modified by the drug release from the dosage form. With an ER dosage form, the apparent absorption takes place in the small intestine, ascending colon, and/or throughout the large intestine depending on the API properties and the product design.[124] A longer absorption phase over an extended period of time, in principle, renders it possible to develop Level A, B, C or Multiple Level C IVIVC. However, matching in vivo drug release remains challenging due to significant heterogeneity in permeability and lumenal environments (eg, pH, water, surface area and contents, transporters, and gut wall metabolism) in different segments of the digestive tract. Thus, the feasibility of an IVIVC remains dependent on the drug molecule, delivery technology, formulation design of the dosage forms, and their interactions with an in vitro test method and condition. Investigation of IVIVC for ER dosage forms should start with understanding drug properties, doses, and their relationship with the associated formulation technologies discussed in chapter "Rational Design of Oral Modified-Release Drug Delivery Systems." In general, the in vivo apparent absorption of ER products containing soluble compounds (eg, BCS class 1) is primarily limited by the drug release from a diffusion or osmotically controlled ER system with fewer significant confounding factors. Thus, an IVIVC is more likely. With BCS class 2 drugs, solubility-to-dose ratio, drug-release mechanism, release duration, and absorption pathways can significantly influence the in vitro and in vivo relationship. In addition, drug particle dissolution following its release (metering) from the dosage form may also play a significant role depending on the dose and solubility of the drug substance. A detailed discussion is provided in an example in the case study section of this chapter. IVIVC for a typical BCS class 3 or 4 drug is more difficult to obtain because of the various factors often involved in its apparent absorption from the GI tract, including low and variable membrane permeability, transporters, and regional dependency.[75,76,78,79,125] In fact, many of the drugs in BCS class 3 or 4 are either not feasible for ER development or exhibit very limited absorption windows,[76,78] such as ranitidine, atenolol, furosemide, erythromycin, cefazolin, amoxicillin, hydrochlorothiazide, methotrexate, acyclovir, and neomycin.

Dosage form behavior as related to IVIVC development depends on drug property, delivery technology, and formulation design. During the drug release from ER systems across different segments of the GI tract, the drug and dosage form are subject to a wide range of environments and conditions, such as varying surface area, absorption pathways, permeability, metabolism, mixing, secretion, lumen content, and amount of fluids. As a result, the in vivo absorption from ER may significantly vary with regions, making its estimation for IVIVC modeling unreliable due to deviations from the system definition of Case III in Table 16.1. In addition, in vitro and in vivo mechanisms may differ or change depending on the drug and formulation design. Increased variability of the

TABLE 16.5 Example of Extended-Release Systems and Food Effect on Bioavailability Parameters

Product	Dosage form	C_{max}	AUC	T_{max}
Theophylline	IR tablet	Decrease (30–50%)	No change	Increase
Theo-Dur	Coated beads	Decrease (60%)	Decrease (50%)	Increase
Theo-24	Coated beads	Increase (120%)	Increase (60%)	No change
Theo-Dur	Matrix tablet	No change	No change	No change
Uniphyl	Matrix tablet	Increase (70%)	Increase (70%)	Increase
Verapamil	IR tablet	Decrease (15%)	No change	Increase
Isoptin SR	Matrix tablet	Decrease (50%)	Decrease (50%)	Increase
Verelan PM	Coated beads	No change	No change	No change
Test Formulation 1[a]	Coated tablet	No change	No change	Increase
Test Formulation 2[a]	Coated beads	No change	No change	Increase
Covera HS	OROS tablet	No change	No change	No change

[a]*Clin Pharmacol Ther, July, 77–83, 1985.*

apparent absorption often observed with ER products may also be exacerbated by improper or less robust formulation designs. A review of literature indicates a wide range of in vivo performance and/or relationships with in vitro drug release for many drugs formulated using different or the same types of delivery technologies. For example, part of the in vivo behavior of a solid product can be inferred from the effect of food on its PK characteristics, that is, food effect. Food intake is known to directly and indirectly induce changes in the GI environment and conditions, including gastric emptying, intestinal motility, mixing, mechanical and shear stress, lumen content, pH, viscosity, ionic strength, osmolality, secretions (bile salts and digestive enzymes), and the activity or capacity of metabolic enzymes and transporters.[27] There is no general in vitro or animal model that is predictive of the effect of such changes on drug absorption. However, the potential impact of these changes on the in vivo performance may be used as an indirect gauge for evaluation of the robustness of an ER solid product in IVIVC development. A survey of food effect on AUC and C_{max} of 47 selected marketed ER products[126] shows that out of 32 soluble APIs, seven exhibit food effect on AUC and eight on C_{max}, while among 15 insoluble APIs, six exhibit food effect on AUC and 14 on C_{max}. Four out of seven osmotic pump products displayed food effect on AUC or C_{max}. Table 16.5 shows ER products of two well-known BCS class 1 drugs and their food effect.[127] It is evident that both ER technology and formulation play a significant role in how a product performs in the GI tract. Therefore, the opportunity and success of developing an IVIVC of ER dosage forms depends on individual APIs as well as the delivery systems and the formulation design.

TABLE 16.6 Common Extended-Release Systems and IVIVC

ER system	Characteristics
Matrix	• In vitro release sensitive to in vitro test conditions. • In vivo results depend on individual drugs and formulation design. • Hydrophilic matrix: Gel strength and system integrity also affect rate and mechanism of drug release in vivo. • Possible to alter in vitro test condition for obtaining IVIVC.
Reservoir	• In vitro release typically sensitive to in vitro test conditions. • In vivo results depend on individual drugs and formulation design. • Possible to adjust test condition for obtaining IVIVC.
Osmotic pump	• In vitro release generally insensitive to test conditions. • In vivo results depend on individual drugs. • Higher probability to obtain IVIVC. • Lack of flexibility to adjust test condition to match in vivo performance.

Table 16.6 is a comparison of three broad types of ER drug delivery systems discussed in chapter "Rational Design of Oral Modified-Release Drug Delivery Systems" with respect to the characteristics related to IVIVC. Among the three types of ER technologies, it is well known that the drug release from an osmotic system is generally insensitive to in vitro test conditions, thus offering a higher chance of matching the in vivo drug release. However, once the in vitro test fails to predict, it is much more difficult to achieve IVIVC via altering the in vitro drug release of an osmotic system. On the other hand, the drug release from the other

two types of systems (reservoir and matrix) is most often dependent on the drug property and in vitro test methodology and conditions, thus providing greater flexibility and opportunity of adjusting the test variables to match the in vitro data with the in vivo profiles. It should be pointed out that for matrix systems, the drug release rate and mechanism are affected by formulation as well as system strength and integrity. The latter is also a function of formulation design. As a result, different formulations of the same drug molecule may exhibit different sensitivity to the test variables, which, in turn, can influence their respective in vitro/in vivo relationships. Therefore, it is important to evaluate key formulation attributes and variables that influence the behavior of a specific dosage form.

16.4 IVIVC DEVELOPMENT APPROACH

The scientific values and significant patient and operational benefits of exploring and establishing an IVIVC in the development of solid dosage forms have been well recognized.

It is important to create an IVIVC strategy and define a rational and effective approach for integration into the product development lifecycle in the project planning stage prior to initiating any development activities.

16.4.1 General strategy and approach

Exploration of IVIVC may start in the early, mid, or late stage of product development depending on the objective and resources of a development program. For example, an IVIVC can be investigated through review and evaluation of the historical in vitro and in vivo data or initiation of an IVIVC study at a later stage of development. This type of retrospective development strategy is usually either driven by the regulatory expectation or due to a lack of awareness of IVIVC and its value in the rational development of a solid product. To fully realize its potential, IVIVC development should begin at the early stage and continue throughout the formulation development cycle if necessary. A prospective or concurrent development strategy requires an integrated evaluation of a drug substance, ER technology, in vitro test method, and IVIVC feasibility followed by incorporating an IVIVC investigation in the formulation screening and development. When the IVIVC is explored as a part of the first in vivo formulation screening studies or subsequent development studies, it can be used to aid in the development of a project timeline, planning of formulation/process studies, and scale-up activities. If an IVIVC exists, it is advantageous to establish the model and validate it if necessary at the early stage to facilitate and accelerate subsequent product development, thereby saving time and resources associated with certain necessary bioequivalence studies. It also helps set the development and regulatory strategy. If an IVIVC is unlikely, a different product development strategy will have to be defined, for example, by planning bioequivalence studies required to support process scale-up and/or certain formulation changes in the development of an ER dosage form.

In the first stage of developing an ER product, two or more prototype formulations with different in vitro release rates are usually tested in humans in order to identify a formulation with a predefined PK performance (see chapter: Rational Design of Oral Modified-Release Drug Delivery Systems). With properly designed formulations and an in vitro test, the study can offer the first opportunity to explore and develop an IVIVC concurrently. For example, to enhance the chance of success or gain insight into an IVIVR, it is essential to consider building similar robustness in the prototype formulations to ensure a consistent release-control mechanism in vivo. This can be achieved via formulation design and additional testing prior to the in vivo test in humans. The former can be accomplished by considering key formulation variables that are likely to affect in vivo performance such as type and level of the rate-controlling materials, properties and loading of the drug and fillers, operating principle, and size and shape of the delivery system in designing the prototypes. The latter may involve challenging the prototypes through testing the in vitro behaviors (integrity, rank order of drug release) under different conditions (apparatus, medium, type and intensity of mixing, physicochemical and mechanical stress) or studying in an appropriate animal model. To fully exploit the information and potential of such in vivo studies for IVIVC, it is also beneficial to select and remain focused on one type of delivery technology in dosage form development. As discussed in chapter "Rational Design of Oral Modified-Release Drug Delivery Systems," when testing prototype formulations with drug-release profiles that bracket the theoretical target absorption profile in vivo, different outcomes may be obtained depending on the API, formulation design, and in vitro test methodology. Fig. 16.10 illustrates four examples of possible in vitro/in vivo relationships of such a study. Unless a quantitative IVIVC is obtained in the first study (Example I), further investigation is required to pursue an IVIVC depending on the outcome. Generally, if little difference is found in the in vivo performance, the in vitro test is considered overdiscriminating (Example IV). It would still be useful to modify the

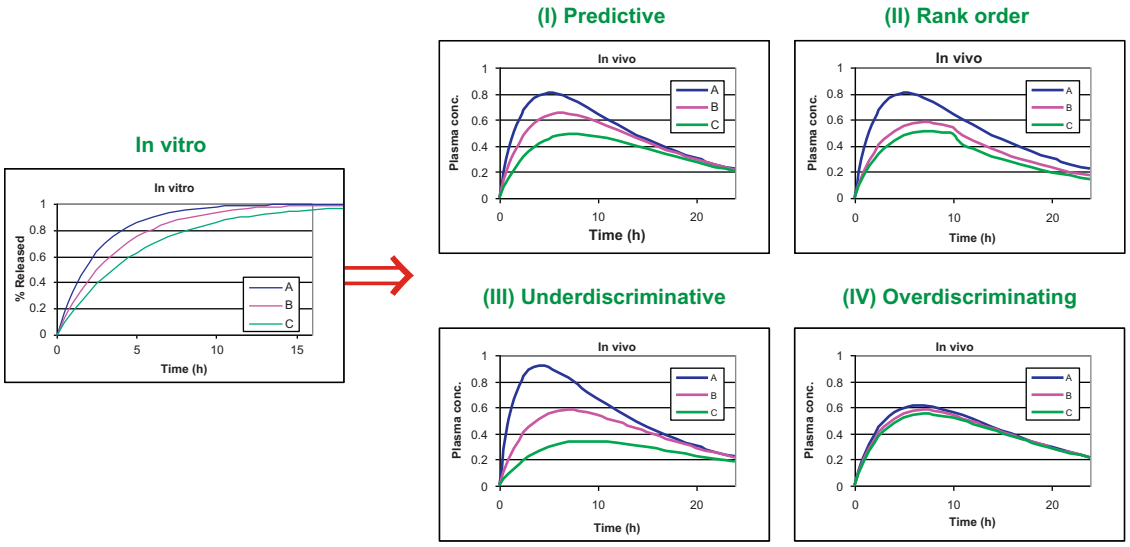


FIGURE 16.10 Examples illustrating possible outcome of the first bioavailability study of prototype ER formulations with varying in vitro release rates.

test conditions to achieve a similar in vitro performance for guiding formulation modifications in the subsequent in vivo studies. If the formulations show in vivo differences (rank order or underdiscriminating), the in vitro test conditions can be modified to correspond with the in vivo data for establishing an IVIVC. Adopting such an IVIVC strategy and making it a part of the dosage form development may take the following sequential steps:

1. Evaluate in vivo behaviors based on the understanding of drug's properties, dosage form, and their interplays with the GI tract.
2. Identify and investigate the mechanism/variables that likely influence or control both in vitro and apparent in vivo drug release.
3. Study test variables to identify condition(s) that:
   a. Differentiates between formulations with different in vivo behaviors.
   b. Deciphers the dosage form behavior and its relationship with the release-controlling variables (eg, critical formulation, processing parameters).
4. Design an in vitro test that matches in vivo performance (if feasible).
5. Establish, evaluate, and validate IVIVC.

It is evident that for formulations with varying PK profiles, developing an IVIVC relies heavily on the successful design of a discriminating in vitro test method.

16.4.2 Design of a predictive in vitro test

The first step in developing a predictive in vitro test is to ensure IVIVC feasibility, that is, that the drug release from the dosage form controls the drug input from the GI tract. While the in vitro data generated for an IVIVC are typically drug release, the estimated in vivo release or apparent absorption is the drug availability in vivo. As discussed previously, the latter is often a function of many other variables in addition to the drug release. Thus, the possibility of using an in vitro release test to predict the in vivo performance is drug and formulation dependent and should be explored on a case-by-case basis.

In the case of IR dosage form of which dissolution is limiting the rate of absorption (eg, certain BCS class 2 or 4 drugs), it is sometimes possible to develop dissolution tests for predicting differences in bioavailability among different formulations and dosing conditions. To achieve an a priori correlation, common approaches include altering the test apparatus or test conditions, some of which are described in Table 16.4. For example, to simulate the composition, volume, and hydrodynamics of the contents in the GI lumen, model compositions of the gastric and intestinal contents before and after meal intake were used as the dissolution media.[105] However, success as measured by meeting the rigorous IVIVC criteria for IR has been scarce, primarily because of the many confounding variables involved in apparent absorption in vivo.

In comparing in vivo with in vitro data of ER dosage forms, it is important to take into considerations a drug's properties, stability, and absorption pathways in relation to the effect of pH, transit time, location of drug release, food, ER technology utilized, and release-control mechanisms.[7] These types of integrated knowledge are essential in anticipating and evaluating the possible interplays of drug-GIT, formulation-GIT and

drug-formulation, and their impacts on the relationship between in vivo and in vivo data. While the results of an in vivo study are invariant, the evaluation of the API characteristics, delivery system, and formulation design can provide greater insight into the in vivo absorption behaviors to help assess IVIVC opportunity and develop a prognostic method.

In general, the observed difference between the in vivo apparent absorption or release mechanism and the in vitro data may be attributed to formulation factors (eg, robustness of the selected delivery technology or a particular composition), test method factors, and/or unique physicochemical and biopharmaceutical properties of the API (eg, low solubility, high dose, enterohepatic recycling, saturable metabolism, truncated absorption, effect of the absorbed drug or metabolite on the GI motility). Some of the drug-related factors can't be simulated by the in vitro test, but they can confound the deconvolution results or make the IR reference inappropriate as the UIR to estimate in vivo absorption if not taken into consideration.[33] For example, the in vivo absorption of the oxybutynin osmotic pump was found to be not only higher than that of the IR reference but also significantly longer (>24 hours) than the in vitro drug release (~15 hours) and the average GI residence time of solid dosage forms as well, resulting in a lack of IVIVC.[128] The former could be attributed to the reduced gut metabolism of the drug released in the lower bowel.[80] However, the latter observation was not understood because oxybutynin has neither known pharmacological effects on GI motility, such as opioid analgesics,[129] nor enterohepatic recycling. Wang et al. compared in vitro and in vivo data of a pulsatile release dosage form of methylphenidate (Ritalin LA). The double peaks were observed in plasma concentration-time profile results from two separate IR doses that were 4 hours apart.[130] This is likely a result of the favorable absorption properties combined with the low dose of methylphenidate. Slower absorption observed for the second phase can be attributed to the decreased in vivo dissolution rate of the second IR due to the changing drug-release environment in the lower bowel.

Sako et al.[131] investigated the effect of composition of a hydrophilic matrix on the in vivo release of acetaminophen. Three formulations containing different fillers showed similar in vitro dissolution profiles using different agitation intensity but had considerable different in vivo performances (Fig. 16.11). The observed discrepancy was believed to result from the different gel strengths of the matrix systems. In developing an in vitro test that differentiates the formulation behaviors, the test method was modified by subjecting the tablets to mechanical stress after 1 hour of dissolution (shaking with glass beads for 10 minutes), which resulted in drug-release profiles similar to those in vivo. It should be mentioned that in addition to creating artificial shear and attrition,[13,131] other techniques that are useful to characterize gel microstructure, formation, and strength of a hydrophilic matrix system include magnetic resonance imaging (MRI) microscopy, confocal microscopy, rheometry, texture analyzer, dynamic mechanical analyzer (DMA), and increasing the ionic strength in the dissolution medium.[132–138]

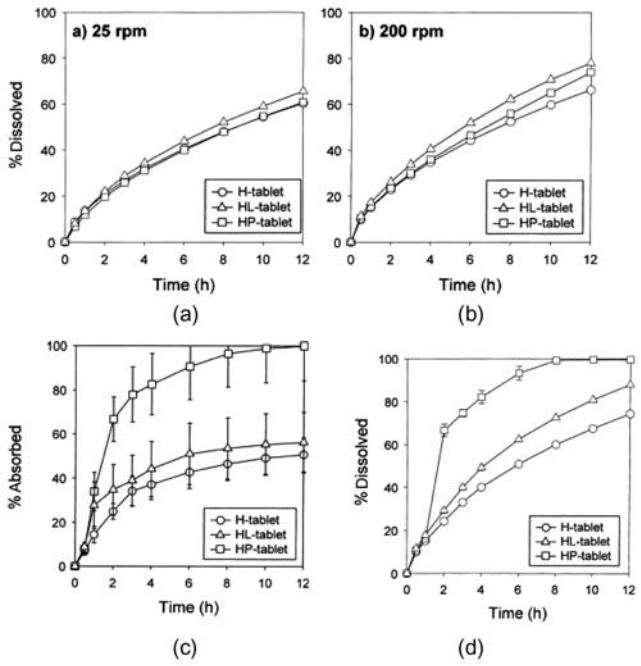


FIGURE 16.11 In vitro drug release and in vivo absorption profiles of three hydrophilic matrix formulations of acetaminophen: (a) and (b) conventional USP test; (c) in vivo absorption; (d) modified USP test. *Reprint with permission: J Control Release 2002;81(1–2):165–172.*

For high dose-to-solubility ratio drug substances that undergo drug dissolution upon delivery (metering) from the ER device, changing the regional dissolution and absorption environment (eg, surface area, amount of water, motility, enzymes, and transporters) may become a significant factor affecting the apparent in vivo absorption, thus IVIVC. For example, a discrepancy between the in vitro and in vivo data of the nifedipine osmotic pump was believed to be a result of the inability of conventional in vitro tests to separate release (metering) from dissolution due to large volume of test medium used throughout testing.[139] Generally, the in vitro test can become overdiscriminative or underdiscriminative of the in vivo results when particle dissolution upon release from an ER system plays a role in absorption, especially if a significant portion of the dose is delivered in the lower bowel.

An additional challenge in exploring a predictive in vitro test for an insoluble compound such as nifedipine is that its extent of absorption also depends on the particle size (or dissolution rate) of the drug upon its release from ER system,[140] which is difficult to reflect using a conventional test method. Furthermore, the potential influence of the solid phase of an insoluble drug substance on in vitro dissolution and in vivo absorption may also complicate the in vitro and in vivo relationship. For example, solid phase transition that occurs in situ or results from precipitation during drug release[141,142] may impact the in vitro and in vivo drug release in a different manner. It should be noted that these types of differences may affect the validity of using an IR reference as the UIR to estimate in vivo absorption.[33]

In investigating IVIVC for dosage forms with more complex release profiles or containing more than one drug substance (eg, bimodal, pulsatile release, and fixed-dose combination), it is essential to take into consideration both the properties of individual actives and their release characteristics. Depending on the release design feature, separate test methods may sometimes be required to match the in vivo absorption of the individual drugs or one part of the profile. It is also possible that a predictive test can only be developed for one of the actives or a portion of the dose and release profiles. For example, for a bimodal release profile consisting of an IR followed by an ER or an ER followed by an IR delivery (Fig. 16.12), it may be feasible to develop an IVIVC for only one (ER) or both portions of the curve. However, special attention is required when developing IVIVC models for more complex MR dosage forms. In a recent paper, Qiu et al. used three model drugs to illustrate the importance of understanding drug properties and product design in establishing and evaluating IVIVC of bimodal MR products.[143] Directly modeling in vitro and in vivo data without first considering drug properties, design characteristics, release mechanisms involved, and product behaviors may lead to a misleading, erroneous outcome or incorrect conclusions in the development of an IVIVC. For products containing more than one drug with synchronized or divergent-release profiles, various IVIVC outcomes are possible depending on a drug's BCS class, biological properties, dose, and IR-to-ER ratio. In any case, a predictive test method will always be useful whether it is for one or both drugs, a portion or the entire absorption profile in setting dissolution specifications and guiding or justifying changes to part of the formulation and process of interest.

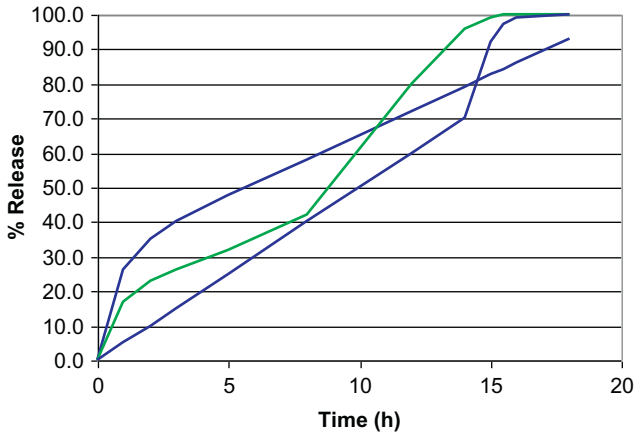


FIGURE 16.12 Illustration of bimodal drug release profiles.

In the method development process, the challenge frequently encountered is that the in vitro release not only differs from the in vivo release/absorption in rate and extent, but it may also exhibit a different release mechanism when a routine pharmacopoeial method is used (eg, USP II, 75 rpm, SIF). A general approach to developing an in vitro test for IVIVC requires identification of the mechanism that controls the in vivo drug release[7,144] For delivery systems that are sensitive to a release environment, the effects of test variables on release kinetics need to be investigated in order to understand how formulation variables and dosage form behaviors respond to environmental changes. In certain cases, formulations can be subject to more stressed conditions (eg, shear stress, high osmolality, ionic strength or agitation intensity, multiple apparatuses) to test robustness or to show difference. Subsequently, a qualitative or quantitative relationship between the drug release and key test variables may be established to guide the selection of a test condition that allows the in vitro data to match both the mechanism and the rates of in vivo release. To define this type of relationship, a well-planned experimental design based on understanding of the API, formulation, and their interplays with in vitro tests is more important than an attempt to simulate particular aspects of GI conditions without considering the specific issues associated with the formulation being studied. For example, switching to FaSSIF and/or FeSSIF media or a multicompartment dissolution system[118,145,146] is unlikely to resolve mismatch problems between in vitro and in vivo data related to a release mechanism change, gel strength, API-polymer interaction in vitro, impact of gut metabolism, or transporters or API on intestinal motility, as discussed previously.

Once an in vitro method that correlates with the in vivo absorption is established, its value as a quality control tool of a drug product is significantly enhanced. It serves as a tool to distinguish between acceptable and unacceptable drug products with respect to the in vivo performance. Lastly, it should be noted that a newly developed method cannot be

considered reliable and acceptable unless it has been validated and demonstrated to be predictive of different formulations with varying in vivo performance.

16.5 APPLICATIONS AND LIMITATIONS

The most important aspect of developing an IVIVC is to identify an in vitro test that can serve as an indicator of how a formulation will perform in vivo and how changes in formulation and/or processing may influence product performance. According to the FDA guidance on IVIVC for extended-release solid dosage forms,[2] a validated IVIVC can be used (1) for setting meaningful dissolution specifications to ensure product quality and in vivo performance and (2) for requesting a biowaiver so that the regulatory burdens, cost, and time associated with product development or postapproval changes can be significantly reduced.

16.5.1 Setting dissolution specifications

In vitro dissolution specifications are established to ensure batch-to-batch consistency and to differentiate between acceptable and unacceptable drug products, thus minimizing the possibility of releasing lots that might not have the desired in vivo performance. In general, dissolution behaviors of the clinical bioavailability batches are used to define the amount released at each time point.[1,2] Dissolution specification settings for IR products should be based on consideration of BCS classification and its ability discriminate with respect to changes in critical material, formulation, and process variables.[1,147] For ER products, challenges often arise in determining the acceptable variation around each time point for ER dosage forms. For new drug applications (NDAs) or abbreviated new drug applications (ANDAs), the specifications are based on the pivotal clinical batch or biobatch of a drug product. In the absence of an IVIVC, the range at any dissolution time point specification has to be within $\pm 10\%$ of the mean profile obtained from the biobatch. A deviation greater than 20% would be acceptable provided that the batches at the specification limits are bioequivalent.[1,2] For ER products, a minimum of three time points covering early, middle, and late stages of the profile are required, with dissolution of at least 80% at the last time point.

In setting dissolution specifications, a validated Level A IVIVC is most useful in ensuring that all lots within the specification limits are bioequivalent. In general, the convolution approach, as illustrated in Fig. 16.9, is preferred, and the specifications should be set on mean data using at least 12 dosage units. In determining the release limits, the dissolution curves defined by the upper and lower extremes established from the biobatch are convoluted to project the corresponding in vivo plasma concentration profiles. A maximum difference of 20% in the predicted C_{max} and AUC is allowed between lots with the fastest and slowest release rates.[2,148] Alternatively, an acceptable set of plasma profiles representing formulations with faster and slower release rates relative to the biobatch can be used to set dissolution specifications by deconvolution based on the principles illustrated in Fig. 16.3. These curves selected based on extremes of 20% difference in C_{max} and AUC are deconvoluted, and the resulting input curves are used to establish the upper and lower dissolution specification ranges at each time point via the IVIVC model.

In the case of Level C and Multiple Level C IVIVC, the specification ranges should be set at the correlation time point such that there is a maximum of 20% difference in the predicted AUC or C_{max}.[2] If the correlation involves more than one parameter, the one resulting in tighter limits should be used. In addition, drug release at the last time point should be at least 80%. Lake et al. reported an example of applying Level C IVIVC to set meaningful dissolution specifications using four carbamazepine IR tablets.[149] In cases where an IVIVC is absent or cannot pass validation criteria, it is still possible, though less optimally, to set a biorelevant specification using lots representing the upper and lower dissolution limits that have been shown to be bioequivalent to the clinical/bioavailability lots or to an appropriate reference standard. An example of establishing this type of biorelevant specification is provided in a case study of this chapter.

Although the general procedure and criteria for the establishment of dissolution specifications that ensure bioequivalence based on IVIVC have been proposed,[2,150,151] few discussions have centered on the detailed process and practical considerations in setting meaningful and realistic specification limits in product development. Elkoshi[152] described a procedure based on release rates that confines C_{max} and AUC values within any desired range to set the minimum range specifications for both zero-order and first-order release products. In reviewing the methods for setting dissolution specifications, Hayes et al.[153] evaluated the most commonly adopted deterministic interpretation approach; that is, those batches passing the in vitro specifications would be bioequivalent, and those failing the specifications would not be bioequivalent if tested in vivo. According to the authors, the deterministic interpretation may not be appropriate, and the conditional probability needs to be considered due to random variation. Through a computer simulation based on an IVIVC model, the conditional

probabilities are shown to depend on the choice of dissolution specifications. The authors further described a method for optimizing the dissolution specifications that take production into consideration. A practical procedure of using IVIVC to establish dissolution specification based on scientific, regulatory, and operation considerations is provided in the case study section of this chapter.

16.5.2 Supporting waiver of in vivo bioavailability study

In addition to serving as a quality control test, comparative dissolution tests have been used to waive bioequivalence or bioavailability studies required for both IR and MR solid dosage forms under certain circumstances. Regulatory requirement including scientific basis, approaches, and evaluation criteria associated with these biowaivers have been clearly laid out in various regulatory guidelines. According to the regulatory guidelines,[1,154–156] the biowaiver request for investigational new drugs (INDs), investigational medicinal product dossiers (IMPDs), NDAs, MAs, ANDAs, and postapproval changes of IR solid dosage forms should be based on the consideration of a drug's BCS class,[157] therapeutic index, and potential effect of excipients on bioavailability. The global regulations with respect to biowaivers for IR solid oral products in the USA, the EU, Japan and from the World Health Organization (WHO) were reviewed by Guptaa et al.[6] To stimulate discussion of biowaivers and methods, the focus group on BCS and biowaivers of the International Pharmaceutical Federation (FIP) to date have published scientific data and biowaiver justifications for a total of 45 BCS class 1-3 compounds in the Journal of Pharmaceutical Sciences,[158] although many of them have not been accepted by regulatory agencies. These drug substances include: acetaminophen, acetazolamide, acetylsalicylic acid, acyclovir, amitriptyline hydrochloride, amodiaquine, hydrochloride, atenolol, bisoprolol fumarate, chloroquine phosphate, chloroquine sulfate, chloroquine hydrochloride, cimetidine, ciprofloxacin hydrochloride, codeine phosphate, diclofenac potassium, diclofenac sodium, doxycycline hyclate, efavirenz, ethambutol dihydrochloride, fluconazole, furosemide, ibuprofen, isoniazid, ketoprofen, lamivudine, levetiracetam, levofloxacin, mefloquine hydrochloride, metoclopramide hydrochloride, metronidazole, nifedipine, piroxicam, prednisolone, prednisone, primaquine diphosphate, propranolol hydrochloride, pyrazinamide, quinidine sulfate, quinine sulfate, ranitidine hydrochloride, ribavirin, rifampicin, stavudine, verapamil hydrochloride, and zidovudine (azidothymidine).

For modified-release products, a dissolution test based on a validated IVIVC can be used for obtaining a waiver for demonstrating in vivo bioavailability often required for NDAs, ANDAs, scale-up, and postapproval changes.[2,159] The criteria for granting the biowaivers using IVIVC are (1) the difference in predicted means of C_{max} and AUC is no more than 20% from that of the reference product, (2) dissolution meets specifications. According to the FDA guidance, categories of biowaivers are also based on the therapeutic index of the drug, the extent of the validation performed on the developed IVIVC, and the dissolution characteristics of the formulation. For instance, for nonnarrow therapeutic index drugs, an IVIVC developed with two formulations can be used for a biowaiver in Level 3 manufacturing site changes and Level 3 nonrelease-controlling excipient changes defined in SUPAC Guidance for MR Solid Dosage Forms.[159] If an IVIVC is developed using three formulations, or two formulations with external validation, a biowaiver may include (1) Level 3 process changes, (2) complete removal or replacement of nonrelease-controlling excipients without affecting the release mechanism, (3) Level 3 changes in the release-controlling excipients, and (4) change of strength (lower than the highest strength).

16.5.3 Limitations and additional considerations

Limitations to the IVIVC methodology reside in the physicochemical, biological, and pharmacokinetic properties of the drug substance and the formulation design, as well as the methodology used to model, evaluate, and validate the IVIVC.

In the development of an IVIVC, the basic assumption of the linear system analysis is that the drug substance exhibits linear pharmacokinetic disposition. Thus, saturable absorption, absorption windows, rate-dependent absorption or rate-dependent presystemic metabolism, and enterohepatic recycling are important factors to consider when modeling and validating an IVIVC because they directly or indirectly result in deviation from the linear assumption.[12,160–163] In addition, an IVIVC should not be developed using plasma concentrations of racemate when there is stereoselective dissolution or absorption between the two enantiomers.[164] More importantly, the dissolution process should be the rate-limiting step in the absorption process, as discussed previously. In most cases, IVIVC models are being established using the average in vivo response, thus ignoring the intersubject and intrasubject variability. For drugs that have relatively high intersubject variability, it is important to take into account the intersubject and intrasubject variability in

constructing and evaluating the IVIVC model.[148,165,166] Cardot and Davit recently described some of the intricacies and possible traps related to the use of mean versus individual data, correction of formulations with different bioavailability, lag time and time scaling, the impact of intersubject and intrasubject variability, and the potential confounding effects of flip-flop kinetics. All of these factors must be considered to increase the chance of a successful IVIVC.[25] Lastly, the in vivo studies used for developing an IVIVC are conducted in healthy volunteers under a well-controlled environment. Factors that might affect the in vivo performance of the dosage form or physiology should also be considered,[167,168] such as food, disease state, age (pediatric and geriatric), and drug–drug or drug-GIT interactions, all of which can affect the GI motility and/or the GI transit time.

The state-of-the-art is such that an IVIVC is typically only valid for one particular type of dosage form containing rate-controlling excipients with the same release mechanism. Even with the same type of solid dosage form, such as a tablet, different release mechanisms (eg, diffusion vs osmosis) often necessitate the development of a separate IVIVC for the same drug molecule. In IVIVC modeling, the absorption parameters obtained with the most widely used Wagner–Nelson method reflect only the rate and not the extent of absorption. Problems can arise from a correlation established using formulations that have different systemic bioavailability. For example, decreased or truncated absorption in the lower GI tract may occur with slow-releasing formulations due to less liquid available for dissolution, lower permeability and surface area, the presence of bacterial metabolism, or a short residence time. As a result, the IVIVC will be apparently formulation-dependent if not corrected. This is illustrated using a simulated example. Two formulations (I and II) were originally designed to release a drug over approximately 8 and 14 hours (Fig. 16.13a). Following oral administration, decreased bioavailability of Formulation II was observed because the window of absorption is found to be approximately 8 hours (Fig. 16.13b). The apparent in vivo absorption profiles of the two formulations obtained by the Wagner–Nelson method are also shown in Fig. 16.13a. A comparison between the in vitro release and the in vivo absorption indicates a good 1:1 relationship for Formulation I and a significant deviation from the relationship for Formulation II as a result of overestimation of the in vivo absorption. Therefore, in developing an IVIVC, reduced AUC needs to be accounted for, for instance, by using time-dependent function: $X_{\text{in vivo}} = g(t)\ X_{\text{in vitro}}$ where $g(t)$ is a step function for truncated or site-dependent absorption.

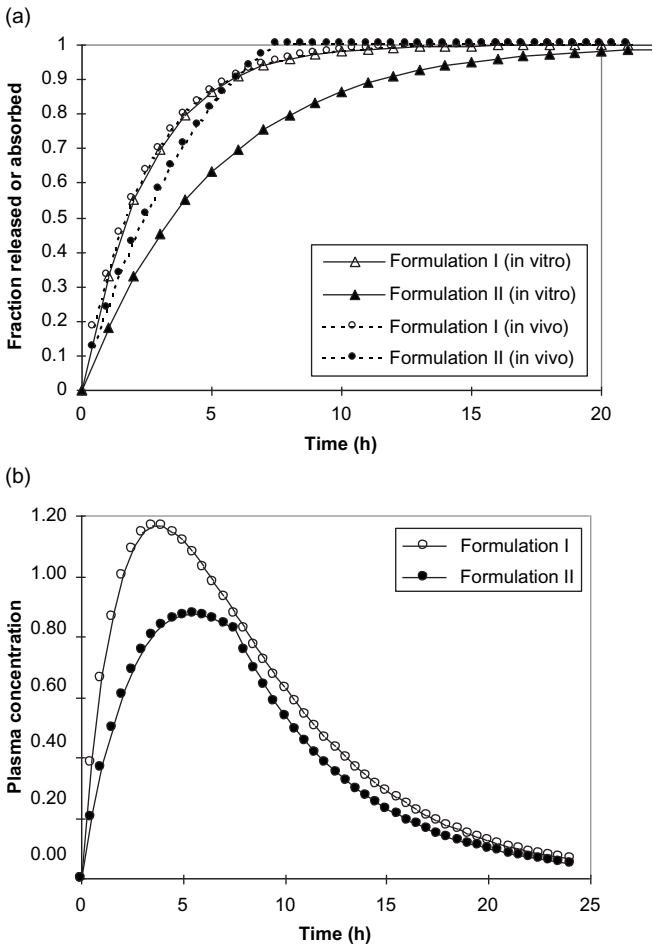


FIGURE 16.13 (a) In vitro release vs in vivo absorption profiles obtained by Wagner–Nelson method based on (b) Simulated plasma concentration profiles of two ER formulations with different release rates.

16.6 CASE STUDIES

16.6.1 Influence of API solubility on IVIVC

Nifedipine is practically insoluble. Its in vivo apparent absorption from the osmotic pump or matrix systems consists of sequential steps of release or metering of drug particles followed by particle dissolution and permeation across the intestinal membrane. This is indirectly supported by the known dependency of bioavailability on the drug particle size discussed previously. Conventional USP tests using a large volume of test medium containing a solubilizer to create sink conditions are incapable of separating the particle dissolution from the drug release. In investigating the IVIVC of Push-Pull osmotic pump of nifedipine, Grundy et al.[103] designed a two-phase test to measure the rate of drug transfer from an aqueous phase into an organic phase, that is, the processes of release of suspension from the device, particle dissolution, and

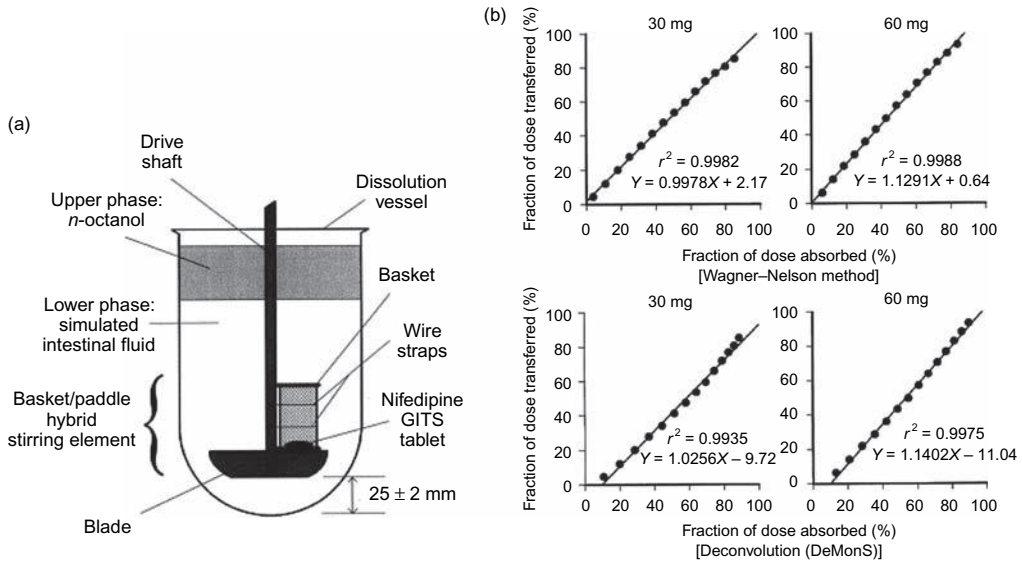


FIGURE 16.14 A Two-phase in vitro test system designed (a) for improving IVIVC of ER dosage form of an insoluble drug, nifedipine (b). *Reprint with permission: J Control Release 1997;48(1–8):9–17.*

subsequent partitioning into the organic phase (Fig. 16.14a). The authors demonstrated that a zero-order rate of drug transfer (0.96 mg/h) obtained from such a test system closely matched the estimated in vivo absorption rate of 1.03 mg/h (30-mg strength) as compared to a rate of 1.7 mg/h based on the conventional test. As a result, an improved 1:1 Level A IVIVC was obtained for all strengths ($R^2 > 0.99$). Similarly, the impact of the drug's solubility on IVIVR was also evaluated in studies comparing ER hydrophilic matrix formulations containing crystalline and amorphous compound with high dose-to-solubility ratio.[85,169] Three tablet formulations containing crystalline API and 10–30% hydroxypropyl methylcellulose (HPMC) exhibited different dissolution rates in a conventional USP method using 900 mL of test medium. However, the in vivo performance of the three formulations is similar, likely a result of a nonsink condition for the in vivo drug release. When the more soluble amorphous drug was used in the same type of ER matrices to improve the particle dissolution, a rank order relationship between the in vitro and in vivo data was observed in the same in vitro test.

16.6.2 Developing a predictive in vitro test[11,12,83]

Depakote ER tablet is a hydrophilic matrix-based, extended-release system with high drug loading. It provides approximately 20 hours of apparent zero-order in vivo absorption (Fig. 16.15). The active ingredient, divalproex sodium, is a stable and permeable compound with pH-dependent solubility. During early

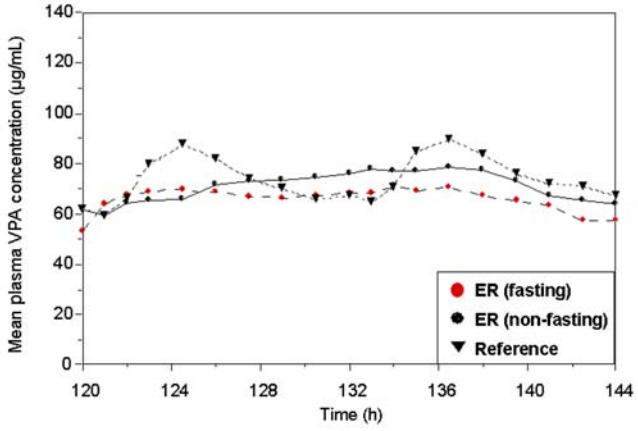


FIGURE 16.15 Mean steady-state plasma concentration profiles of once-daily Depakote ER tablet dosed under fasting and nonfasting conditions with twice-daily enteric Depakote tablet as reference.

formulation development, the in vitro drug release of three different formulations were all found to be slower and showed inadequate separation compared to in vivo absorption (underdiscriminating) when a conventional test was used (Fig. 16.16a). In addition, the mechanism of the in vitro release was diffusion controlled, whereas the apparent absorption profile obtained by deconvolution showed zero-order absorption, suggesting a predominantly erosion-controlled in vivo release (also supported by steady-state plasma concentration curves in Fig. 16.15).

In order to develop a new in vitro test that predicts in vivo absorption, statistically designed studies were carried out to investigate the effects of various in vitro testing variables on drug release. The variables

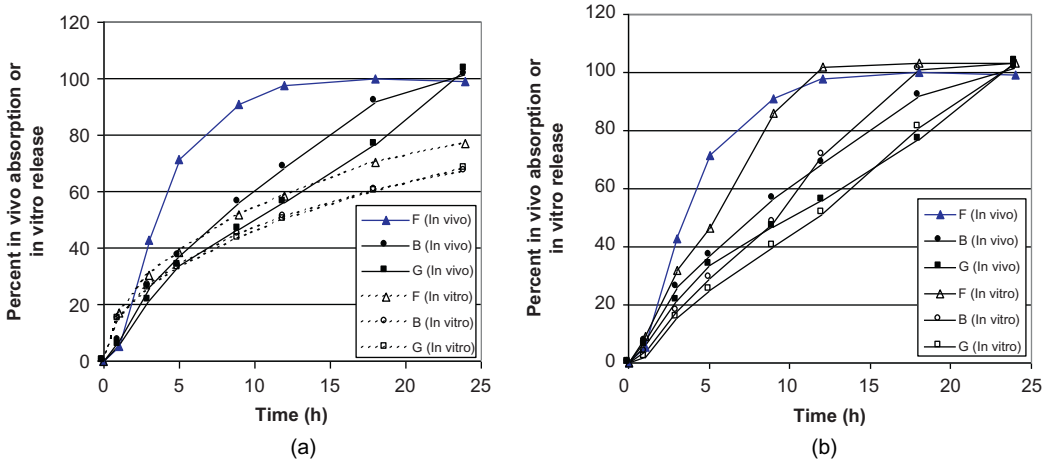


FIGURE 16.16 Mean in vivo absorption versus in vitro release profiles of three formulations using (a) conventional in vitro test method (USP II, 100 rpm, pH 7.5 phosphate buffer) and (b) a predictive test method (USP II, 100 rpm, 0.75 h 0.1 N HCl followed by 0.05 M phosphate buffer containing 75 mM SLS at pH 5.5).

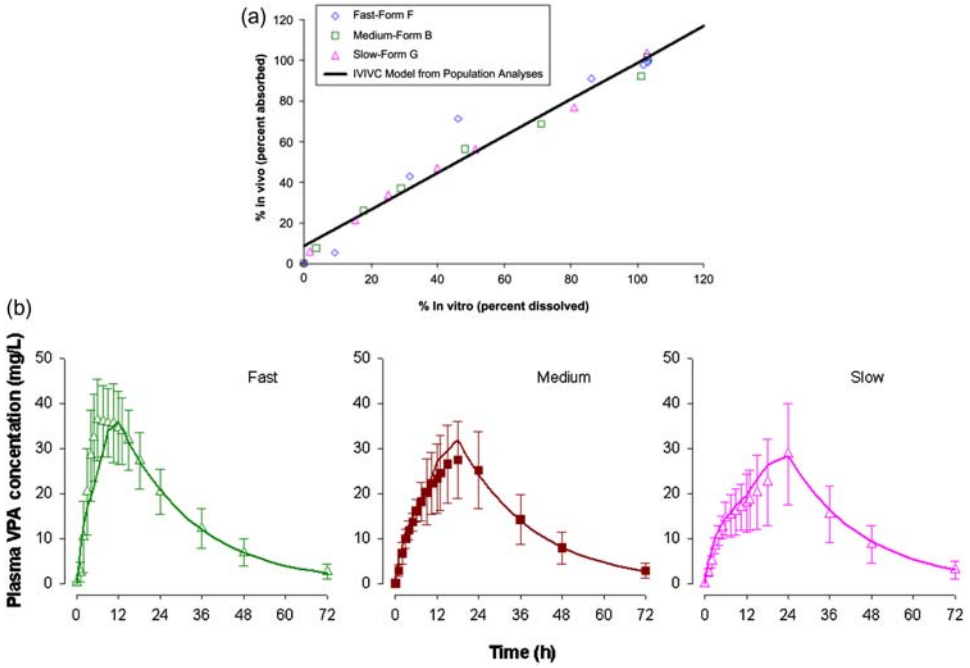


FIGURE 16.17 Results of in vivo/in vitro correlation studies of extended-release divalproex sodium tablets: (a) IVIVC plot and (b) mean predicted in vivo plasma profiles of three different formulations.

investigated included agitation intensity, apparatus, surfactant, pH, and ionic strength of the dissolution medium. Based on factorial studies and statistical analysis, a new set of test conditions was determined and demonstrated to correlate with the in vivo drug absorption for various ER formulations (Fig. 16.16b). Statistical evaluation of the in vitro method based on a hypothesis test indicated that the same IVIVC equation holds for the three different formulations (Fig. 16.17a). A mixed-effects model was used for data analysis in which the dependence among observations from the same subject in the human pharmacokinetic study was taken into account. Fig. 16.17b shows the agreement between the observed and the predicted plasma profiles of the tablets. The method has been successfully validated internally and externally on multiple occasions over a period of eight years. More importantly, it has been applied to (1) set meaningful drug-release specifications, (2) justify biowaivers for postapproval changes, and (3) assure product quality by timely

detecting significant dissolution changes during commercial production due to a change of the key rate-controlling polymer property that could not have been detected had the conventional dissolution test been used as a QC tool.[170]

This study illustrates a useful approach in identifying a predictive method for the development of an IVIVC, that is, adjusting the dissolution test conditions to correlate the in vitro data with the in vivo behaviors of the formulations.

16.6.3 Illustration of setting an optimal dissolution specification based on IVIVC using Monte Carlo simulation

The ability to establish a meaningful dissolution specification acceptable for commercial manufacturing has profound implications in ensuring consistent product performance and the routine production of the solid products. A specification based on product and process understanding assures both product quality and supply. If the specification range is unnecessarily narrow, the probability of failing a batch is increased due to the inherent variability of the raw materials, the product, and the manufacturing process. If a wide range is set for passing batches or solely based on the individual product and process capability, the in vivo performance may not always be ensured, especially in the absence of an IVIVC. One of the most significant advantages of establishing dissolution acceptance criteria using an IVIVC is that it offers greater flexibility for identifying a specification that maximize the probabilities of an assured product in vivo performance and a successful commercial production. More specifically, it allows for searching an optimal range in a multidimensional space defined by the needs of bioequivalence, quality control, manufacturability, and regulation.

One of major challenges in arriving at an optimal dissolution specification prior to regulatory filing is a lack of sufficient data to measure process capability at full production scale. To address this problem, Monte Carlo simulation can be used to evaluate acceptability of the proposed dissolution specifications with respect to manufacturability, bioequivalence, and regulatory requirement based on the inherent material, product, and process variability.[171–173] Monte Carlo simulation is a statistical tool for stochastic model calculations and analysis of propagation from uncertainties in model inputs into uncertainties in outputs (results).[174] It is commonly used to assess risk by making use of a pseudorandom drawing that simulates the real-life sampling to produce distributions of different outcome values. Its core idea is to use probability curves and random samples of inputs to explore the behavior of a complex stochastic system and determine/model the probability of different outcomes that cannot readily be predicted due to the intervention of random variables. Monte Carlo simulation generally includes four basic steps[175]: (1) define a domain of possible inputs, (2) generate inputs randomly from a probability distribution over the domain, (3) perform a deterministic computation on the inputs, and (4) aggregate the outputs for analysis. A number of statistical software, such as Minitab, Matlab, and MS Excel are often used to run Monte Carlo simulations.

This case study describes a useful approach to identify the dissolution specification limits that ensure product quality, performance, and robustness of commercial manufacturing. It is a data-driven approach that utilizes the available dissolution data, a validated IVIVC model, and Monte Carlo simulation, consisting of the following steps:

1. Based on the pivotal biobatch, generate multiple sets of specification ranges that meet bioequivalence criteria by confining differences in IVIVC model-predicted C_{max} and AUC values within 20%.
2. Review and gather all existing dissolution data, which often include individual values of hundreds or thousands of individual tablets generated from a pilot scale, a larger scale, and stability studies during product development, and perform statistical analysis to estimate the probability distribution and variability of dissolution data resulting from natural variations in raw materials, product, processing, and test method.
3. Perform Monte Carlo simulation to select an acceptable specification that accommodates the inherent system variability. Specifically, dissolution data are generated using Monte Carlo simulation based on the statistical distribution (mean and spread) of the existing representative data. The simulated data are tested against the proposed specification limits using dissolution stage testing criteria (L1, L2, and L3). The simulation is typically iterated tens of thousands of times to estimate the overall probability of failing different stage testing associated with each set of proposed specification. For example, 100,000 production lots can be simulated by randomly sampling 100,000 groups of six tablets for dissolution stage testing when assessing individual sets of specification. The simulation is repeated for every set of proposed specifications until a set of specification is identified with minimum risk of failing the dissolution stage testing while ensuring all batches within the lower and upper specification limits are bioequivalent to one another.

Table 16.7 shows an illustrative example of this approach. Monte Carlo simulations were conducted

TABLE 16.7 Monte Carlo Simulation Results for Assessing Four Sets of Dissolution Specifications of an Extended-Release Tablet Based on IVIVC and Manufacturing Consideration

Specification set	Q (2 h)	Q (8 h)	Q (12 h)	P[a] (Fail L3)	P (pass L1)	P (pass L2)	P (pass L3)
I	NML 20%	35–66%	NLT 80%	0.0%	89.3%	10.2%	0.5%
II	NML 18%	33–62%	NLT 80%	0.0%	78.2%	12.3%	0.5%
III	NML 18%	36–60%	NLT 80%	0.1%	75.4%	24.4%	0.1%
IV	NML 20%	38–58%	NLT 80%	0.2%	72.9%	26.1%	0.8%

[a]*P: Probability of passing or failing dissolution stage testing.*

using four sets of specifications that meet the in vivo bioequivalence criteria for an ER tablet product. Sets I and II are considered acceptable for commercial manufacturing because they show a minimum risk of batch failures. It is worth noting that the accuracy and the reliability of this approach depend on the quality and the representativeness of the database used to estimate the inherent variability. Data from the commercial production following product approval should be collected and used to verify the simulation results and further improve the model if necessary.

16.6.4 Setting clinically relevant specifications

In science-based pharmaceutical development, IVIVC is recognized as one of the most important biopharmaceutical tools for enhancing drug product and process understanding and for ensuring consistent efficacy and safety throughout the drug product's lifecycle because it allows for the establishment of clinically relevant product specifications. Clinically relevant product specifications may be defined as those that can ensure the delivery of the intended dose at a consistent rate to patients to guarantee safety and efficacy profiles for the marked product relative to those achieved by the clinical trial formulations. By linking product quality to the clinical performance in product development, manufacturing, and continual improvement, clinically relevant specifications or controls associated with product critical quality attributes (CQA), critical material attributes (CMA), and critical process parameters (CPP) are expected to assure high product quality with a consistent safety and efficacy profile desired throughout lifecycle of a drug product. As a bridge between in vitro testing and in vivo exposure, an IVIVC can be effectively employed for understanding the impact of the drug product CQAs on in vivo performance. Through IVIVC, product quality specifications can be established optimally with assured clinical outcomes, using in vitro dissolution testing as a surrogate. For example, Duan et al. reported the use of IVIVC in setting the clinical relevant specifications for an ER matrix tablet.[176] During the development, three CMAs were identified as high-risk factors: particle size distribution (PS) of the API, viscosity of a matrix forming agent (MFA), and the PS of a matrix forming enhancing agent (MEA). Their impact on dissolution was investigated using a design of experiments (DOE). The data were analyzed by partial least square regressions and validated by comparing the model predicted and the observed dissolution values, along with a "leave one out" cross validation. The validated model was first used to predict dissolution at different values of MFA viscosity, particle sizes of MEA, and the API. The corresponding effects on AUC and C_{max} were subsequently evaluated using a multiple Level C IVIVC. The predicted AUCs and C_{max}s were compared to those of the clinical batches with demonstrated efficacy and safety, and the ratios of AUC and C_{max} were calculated and plotted against the three CMAs. For instance, Fig. 16.18 shows the combined effect of MFA viscosity and MEA PS on AUC ratio predicted using dissolution data at 3 hours under four different values of API particle size, expressed as the ratios of larger to smaller particle size (3.2, 6.6, 12.1 and 14.4). The rectangular box encloses the acceptable spaces for setting appropriate specifications, since they fall well within the bioequivalence range (0.8–1.25).

16.6.5 Setting biorelevant dissolution specification

Methylphenidate (MPH) is an amphetamine-like central nervous system stimulant commonly prescribed to treat attention-deficit/hyperactivity disorder (ADHD) in children, adolescents, and adults, as well as narcolepsy. It is a weak base with a pK_a of 8.77 and a log$P$ of 3.19. Its hydrochloride salt is freely soluble in water (18.6 mg/mL), stable, and well absorbed from the intestinal tract with a short elimination half-life of 3–4 hours.[177] These favorable properties combined with a low dose make MPH an ideal candidate for oral MR delivery. As a result, products using different MR

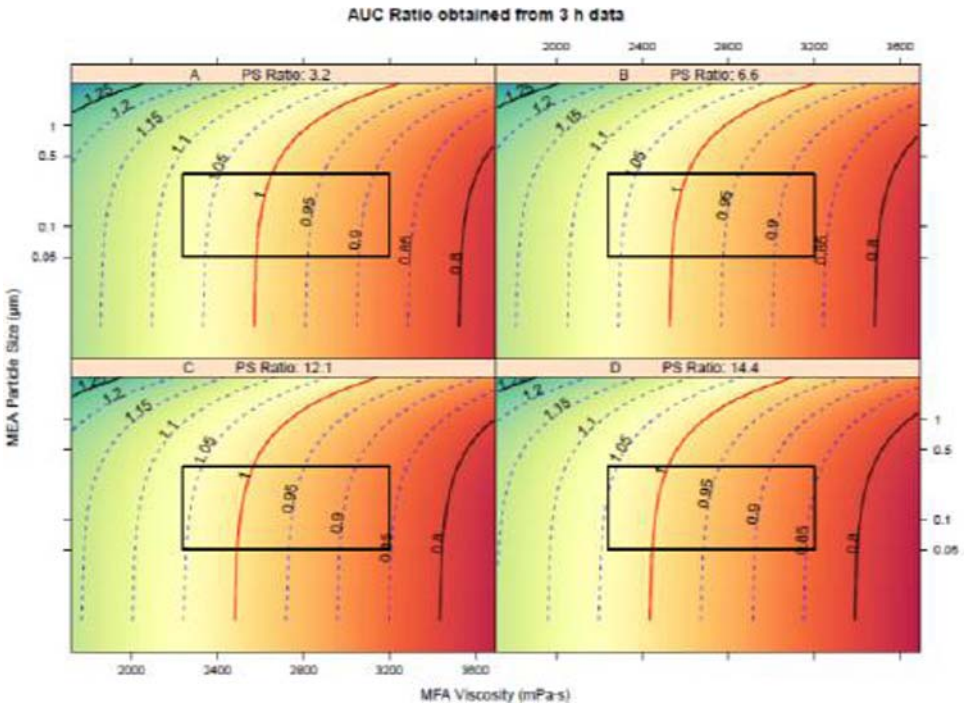


FIGURE 16.18 Effect of different combinations of three CMAs on AUC ratios predicted using the in vitro data at 3 h based on a Multiple Level C IVIVC.

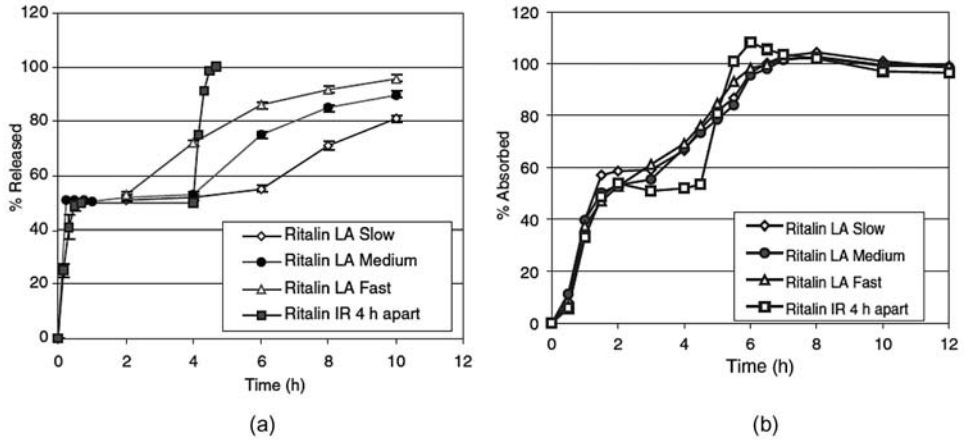


FIGURE 16.19 In vitro dissolution and in vivo absorption profiles of three Ritalin LA formulations and Ritalin IR tablet administered 4 hours apart.

design principles have been commercialized (see chapter: Product and Process Development of Solid Oral Dosage Forms). Wang et al. reported a study that evaluated the in vitro dissolution and in vivo absorption of MPH from a pulsatile release product (Ritalin LA capsule) using an IR formulation (Ritalin IR tablet) as a reference.[130] The Ritalin LA capsule consisting of 50% IR and 50% DR beads was designed to mimic PK performance of the IR product given 4 hours apart. In the study, three formulations, slow-, medium-, and fast-release, were prepared by varying the release rate of the DR portion by coating the IR beads to a different weight gain with an acrylates copolymer (Fig. 16.19a). The in vitro dissolution of the three MR formulations was determined using USP apparatus 1 (100 rpm) in 0.01N HCl for 2 hours followed by a pH 6.8 phosphate buffer. The three formulations were evaluated in 18 healthy volunteers under fasted conditions using a single-dose, randomized, four-way crossover design.

It was found that the three test formulations exhibited similar plasma concentration-time profiles with two peak concentrations and were bioequivalent

with respect to C_{max} and AUC and the corresponding values describing the first and second peaks, although the in vitro dissolutions were different.[130] Fig. 16.19b showed that the absorption of MPH was biphasic, with a rapid absorption phase between 0 and ~2 hours and a slightly slower second absorption between ~3 and 6 hours. The results of this study can be used to justify and establish the biorelevant dissolution specification by defining a dissolution space that ensures bioequivalence in the absence of an IVIVC. More specifically, a wide range of dissolution specifications can be set based on the in vitro profiles of the fast and slow formulations, as shown in Fig. 16.19a, because batches prepared within the limits have been shown to be bioequivalent.

16.7 SUMMARY

The general concepts, theory, modeling methodology, assessment, and applications of in vitro/in vivo correlation have been established and extensively investigated, though differences in scientific approaches remain in the details of model development and evaluation. The state of the art is such that there is no universal in vitro model that can mimic or reproduce the highly complex and dynamic GI environment or predict the in vivo performance of solid oral dosage forms. Therefore, development of an IVIVC must be carried out case by case based on the understanding of the API properties, product characteristics, and their interplay with in vitro test method/conditions.

IVIVC is generally more likely for ER dosage forms than IR products, since drug absorption is typically limited by the drug release. To increase the chance of success, it is crucial to evaluate IVIVC feasibility of in vitro and in vivo results by integrating knowledge of physicochemical and biopharmaceutical characteristics of drug substances, dosage form design, and their interplays with the GI environment and in vitro test conditions. It is also important to make IVIVC strategy an integral part of the dosage form development program.

Once an IVIVC is developed and validated, the predictive in vitro test can be used as a surrogate for in vivo studies, a guide for setting meaningful product specifications, and a reliable tool for quality control. Whenever feasible, an IVIVC-based in vitro test method should be implemented in the QC laboratories such that any potential or unexpected changes of in vivo performance of a product during production can be detected to ensure safety and efficacy of every commercial batch.

References

1. Guidance for industry: dissolution testing of immediate release solid oral dosage forms. U.S. Department of Health, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), August, 1997.
2. Guidance for industry: extended release oral dosage forms: development, evaluation, and application of in vitro/in vivo correlations. U.S. Department of Health, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), September, 1997.
3. Note for guidance on development pharmaceutics; European agency for the evaluation of medicinal products. Human Medicines Evaluation Unit, Committee for Proprietary Medicinal Products, CPMP/QWP/155/96; London, January 1998.
4. Note for guidance on modified release oral and transdermal dosage forms: section II (Pharmacokinetic and Clinical Evaluation). European Agency for the Evaluation of Medicinal Products, Human Medicines Evaluation Unit, Committee for Proprietary Medicinal Products, CPMP/EWP/280/96, London, July 1999.
5. Guideline on the pharmacokinetic and clinical evaluation of modified release dosage forms. European Medicines Agency, Committee for Medicinal Products for Human Use (CHMP), (EMA/CPMP/EWP/280/96 Corr1). London, November 2014.
6. Guptaa E, Barendsb DM, Yamashitaa E, Lentzc KA, Harmszed AM, Shahe VP, et al. Review of global regulations concerning biowaivers for immediate release solid oral dosage forms. *E. J Pharm Sci* 2006;**29**(3-4):315–24.
7. Qiu Y. Design and applications of predictive in vitro tests for in vitro-in vivo correlations of oral extended-release dosage forms. *Am Pharm Rev* 2006;**9**(1):94–9.
8. Souliman S, Blanquet S, Beyssac E, Cardot J-M. A level A in vitro/in vivo correlation in fasted and fed states using different methods: applied to solid immediate release oral dosage form. *Eur J Pharm Sci* 2006;**27**(1):72–9.
9. Rossia RC, Diasa CL, Donatoa EM, Martinsa LA, Bergolda AM, Fröehlich PE. Development and validation of dissolution test for ritonavir soft gelatin capsules based on in vivo data. *Inter J Pharm* 2007;**338**(1-2):119–24.
10. Morita R, Honda R, Takahashi Y. Development of a new dissolution test method for an oral controlled release preparation, the PVA swelling controlled release system (SCRS). *J Control Release* 2003;**90**(1):109–17.
11. Qiu Y, Garren J, Samara E, Abraham C, Cheskin HS, Engh KR. Once-a-day controlled-release dosage form of divalproex sodium II: development of a predictive in vitro release method. *J Pharm Sci* 2003;**92**(11):2317–25.
12. Dutta S, Qiu Y, Samara E, Cao G, Granneman GR. Once-a-day controlled-release dosage form of divalproex sodium III: development and validation of a level a in vitro-in vivo correlation (IVIVC). *J Pharm Sci* 2005;**94**:1949–56.
13. Sirisuth N, Augsburger LL, Eddington ND. Development and validation of a non-linear IVIVC model for a diltiazem extended release formulation. *Biopharm Drug Disposition* 2002;**23**(1):1–8.
14. Tabusso G. Regulatory aspects of development pharmaceutics (2). *Regulatory Affairs J* 1992;**12**:909–12.
15. Pharmacopeial Forum. In vitro and in vivo evaluation of dosage forms 1993;**19**:5366–5379.
16. Skelly JP, et al. Report of the workshop on controlled release dosage forms: issues and controversies. *Pharm Res* 1987;**4**:75–8.
17. Pharmacopeial Forum Stimuli Article. In vitro-in vivo correlation for extended release oral dosage forms 1988;4160–4161.
18. Skelly JP, Shiu GF. In vitro/in vivo correlations in biopharmaceutics: scientific and regulatory implications. *Eur J Drug Metab Pharm* 1993;**18**:121–9.

19. Siewert W. Perspectives of in vitro dissolution tests in establishing in vitro/in vivo correlations. *Eur J Drug Metab Pharm* 1993;**18**:7–18.
20. Skelly JP, et al. In vitro and in vivo testing and correlation for oral controlled/modified-release dosage forms. *Pharm Res* 1990; **7**(9):975–82.
21. Khan MZ. Dissolution testing for sustained or controlled release oral dosage forms and correlation with in vivo data: challenges and opportunities. *Int J Pharm* 1996;**140**:131–43.
22. Gould AL, Agrawal NG, Goel TV, Fitzpatrick S. A 1-step Bayesian predictive approach for evaluating in vitro in vivo correlation (IVIVC). *Biopharm Drug Dispos* 2009;**30**(7):366–88.
23. Mendyk A, Tuszyński PK, Polak S, Jachowicz R. Generalized in vitro-in vivo relationship (IVIVR) model based on artificial neural networks. *Drug Des Devel Ther* 2013;**7**:223–32.
24. Costello C, Rossenu S, Vermeulen A, Cleton A, Dunne A. A time scaling approach to develop an in vitro-in vivo correlation (IVIVC) model using a convolution-based technique. *J Pharmacokinet Pharmacodyn* 2011;**38**(5):519–39.
25. Cardot J-M, Davit BM. In vitro–in vivo correlations: tricks and traps. *AAPS J* 2012;**14**(3):491–9.
26. Benet L. The role of BCS (biopharmaceutics classification system) and BDDCS (biopharmaceutics drug disposition classification system) in drug development. *J Pharm Sci* 2013;**102**(1):34–42.
27. Custodio JM, Wu CY, Benet LZ. Predicting drug disposition, absorption/elimination/transporter interplay and the role of food on drug absorption. *Adv Drug Deliv Rev* 2008;**60**(6):717–33.
28. Darwich AS, Neuhoff S, Jamei M, Rostami-Hodjegan A. Interplay of metabolism and transport in determining oral drug absorption and gut wall metabolism: a simulation assessment using the "Advanced Dissolution, Absorption, Metabolism (ADAM)" model. *Curr Drug Metab* 2010;**11**(9):716–29.
29. Larregieu CA, Benet LZ. Distinguishing between the permeability relationships with absorption and metabolism to improve BCS and BDDCS predictions in early drug discovery. *Mol Pharm* 2014;**11**(4):1335–44.
30. Mojaverian P, et al. *In vivo-in vitro* correlation of four extended release formulations of pseudophedrine sulfate. *J Pharm Biomed Analysis* 1997;**15**:439–45.
31. Mojaverian P, et al. Correlation of *in vitro* release rate and *in vivo* absorption characteristics of four chlorpheniramine maleate extended release formulations. *Pharm Res* 1992;**9**:450–6.
32. Young D, DeVane J, Butler J, editors. *In vitro-in vivo correlations*. New York, NY: Plenum Press; 1997.
33. Qiu Y, Samara E, Cao G. In vitro/in vivo correlations in the development of solid controlled release dosage forms. In: Wise DL, Klibanov AM, Langer R, Mikos AG, Peppas NA, Trantolo DJ, Wnek GE, Yaszeski MJ, editors. *Handbook of pharmaceutical controlled release technology*. New York, NY: Marcel Dekker, Inc; 2000. p. 527–49.
34. Pharmaceutical product development. In: Chilukuri DM, Sunkara G, Young D, editors. *Vitro-in vivo correlation*. New York, NY: Informa Healthcare USA Inc; 2007.
35. Heeger D. Signals, Linear Systems, and Convolution. 2000. <http://www.cns.nyu.edu/~david/handouts/convolution.pdf>.
36. Cutler DJ. Linear system analysis in pharmacokinetics. *J Pharmacokin Biopharm* 1978;**6**:265–82.
37. Moller H. Deconvolution techniques and their use in biopharmaceutics. In: Hardy JG, Davis SS, Wilson CG, editors. *Drug delivery to the gastrointestinal tract*. Chichester: Ellis Horwood Ltd; 1989. pp. 179–94.
38. Cutler DJ. Numerical deconvolution by least squares: use of prescribed input functions. *J Pharmacokin Biopharm* 1978;**6**:227–41.
39. Cutler DJ. Numerical deconvolution by least squares: use of polynomials to represent the input function. *J Pharmacokin Biopharm* 1978;**6**:243–63.
40. Langenbucher F. Numerical convolution/deconvolution as a tool for correlating *in vitro* with *in vivo* drug availability. *Pharm Ins* 1982;**4**:1166–72.
41. Gillespie WR, Veng-Pedersen P. A ployexponential deconvolution method. Evaluation of the "gastrointestinal bioavailability" and mean *in vitro* dissolution time of some ibuprofen dosage forms. *J Pharmacokin Biopharm* 1985;**13**:289–307.
42. Vajda S, Godfrey KR, Valko P. Numerical deconvolution using system identification methods. *J Pharmacokin Biopharm* 1988;**16**:85–107.
43. Verotta D. Two constrained deconvolution methods using spline functions. *J Pharmacokin Biopharm* 1993;**21**:609–36.
44. Vaubhan DP, Dennis M. Mathematical basis of the point area deconvolution method for determining *in vivo* input functions. *J Pharm Sci* 1978;**67**:663–5.
45. Veng-Pedersen P. Model independent method of analyzing input in linear pharmacokinetic systems having polyexponential impulse response. 1: theoretical analysis. *J Pharm Sci* 1980; **69**:298–304.
46. Veng-Pedersen P. Model independent method of analyzing input in linear pharmacokinetic systems having polyexponential impulse response. 2: numerical evaluation. *J Pharm Sci* 1980; **69**:305–12.
47. Veng-Pedersen P. Novel deconvolution method for linear pharmacokinetic systems with polyexponential impulse response. 1: theoretical analysis. *J Pharm Sci* 1980;**69**:312–18.
48. Veng-Pedersen P. Novel approach to bioavailability testing: statistical method for comparing drug input calculated by a least squares deconvolution technique. *J Pharm Sci* 1980;**69**:319–24.
49. Madden FN, Godfrey KR, Chappell MJ, Hovorka R, Bates RA. A comparison of six deconvolution techniques. *J Pharmacokin Biopharm* 1996;**24**:283–99.
50. Veng-Pedersen P. Personal communications.
51. Hanaano M. Studies on absorption and excretion of drug. VII. A new estimation method for the release of drugs from dosage forms and the availability in vivo. *Chem Pharm Bull* 1967;**15**(7):994–1001.
52. Nogami H, Hanano M. Studies on absorption and excretion of drug VII. The release rate in vivo of antipyretic and analgesic drugs. *Chem Pharm Bull* 1967;**15**(7):1002–9.
53. Chiou WL. New compartment- and model-independent method for rapid calculation of drug absorption rates. *J Pharm Sci* 1980;**69**(1):57–62.
54. Langenbucher F. Improved understanding of convolution algorithms correlation body response with drug input. *Pharm Ind* 1982;**44**(12):1275–8.
55. Langenbucher F. Handling of computational in vitro/in vivo correlation problems by Microsoft Excel: III. Convolution and deconvolution. *Eur J Pharm Biopharm* 2003;**56**(3):429–37.
56. Wagner JG. *Absorption analysis and bioavailability*. *Pharmacokinetics for the pharmaceutical scientist*. Lancaster, PA: Technomic Pub. Co. Inc.; 1993. pp. 159–206.
57. Loo J, Riegelman S. New method for calculating the intrinsic absorption rate of drugs. *J Pharm Sci* 1968;**57**:918–28.
58. Wagner JG. Pharmacokinetic absorption plots from oral data alone or oral/intravenous data and an exact Loo-Riegelman equation. *J Pharm Sci* 1983;**72**:838–42.
59. Veng-Pedersen P. Mean time parameters in pharmacokinetics. Definition, computation and clinical implications (Part I). *Clin Pharmacokin* 1989;**17**:345–66.
60. Podczeck F. Comparison of *In vitro* dissolution profiles by calculating mean dissolution time (MDT) or mean residence time (MRT). *Int J Pharm* 1993;**97**:93–100.
61. Gillespie WR. Convolution-based approaches for in vivo-in vitro correlation modeling. *Adv Exp Med Biol* 1997;**423**:53–65.

62. Veng-Pedersen P, Gobburu JVS, Meyer MC, Straughn AB. Carbamazepine level-A in vivo-in vitro correlation (IVIVC): a scaled convolution based predictive approach. *Biopharm Drug Disposition* 2000;**21**(1):1–6.
63. Gillespie WR. *In vivo* Modeling Strategies for IVIVC for Modified Release Dosage Forms. AAPS/CRS/FDA Workshop on Scientific Foundation and Applications for the Biopharmaceutics Classification System and *In vitro-In vivo* Correlations. Arlington, VA, April 1997.
64. Ritger PL, Peppas NA. A simple equation for description of solute release I. Fickian and non-Fickian release from non-swellable devices in the form of slabs, spheres, cylinders or disck. *J Control Release* 1987;**5**:23–6.
65. O'Hara T, Hayes S, Davis J, Devane J, Smart T, Dunne A. In vivo-in vitro correlation (IVIVC) modeling incorporating a convolution step. *J Pharmacokinet Pharmacodyn* 2001;**28**(3):277–98.
66. Gaynor C, Dunne A, Davis J. A comparison of the prediction accuracy of two IVIVC modeling techniques. *J Pharm Sci* 2008;**97**(8):3422–32.
67. Jacobs T, Rossenu S, Dunne A, Molenberghs G, Straetemans R, Bijnens L. Combined models for data from in vitro-in vivo correlation experiments. *J Bioph arm Stat* 2008;**18**(6):1197–211.
68. Veng-Pedersen P, Widness JA, Pereira LM, Peters C, Schmidt RL, Lowe LS. Kinetic evaluation of nonlinear drug elimination by a disposition decomposition analysis. Application to the analysis of the nonlinear elimination kinetics of erythropoietin in adult humans. *J Pharm Sci* 1995;**84**(6):760–7.
69. Mumenthaler MS, Taylor JL, Yesavage JA. Ethanol pharmacokinetics in white women: nonlinear model fitting versus zero-order elimination analyses. *Alcohol Clin Exp Res* 2000;**24**(9):1353–62.
70. Eadie MJ, Tyrer JH, Bochner F, Hooper WD. The elimination of phenytoin in man. *Clin Exp Pharmacol Physiol* 1976;**3**(3):217–24.
71. Gaynor C, Dunne A, Davis J. A comparison of the prediction accuracy of two IVIVC modelling techniques. *J Pharm Sci* 2008;**97**(8):3422–32.
72. Gaynor C, Dunne A, Costello C, Davis J. A population approach to in vitro–in vivo correlation modelling for compounds with nonlinear kinetics. *J Pharmacokinet Pharmacodyn* 2011;**38**:317–32.
73. Gibaldi M, Perrier D. *Pharmacokinetics*. 2nd ed. New York, NY: Dekker; 1982.
74. Kesisoglou F, Hermans A, Neu C, Yee KL, Palcza J, Miller J. Development of in vitro-in vivo correlation for amorphous solid dispersion immediate-release suvorexant tablets and application to clinically relevant dissolution specifications and in-process controls. *J Pharm Sci* 2015;**104**(9):2913–22.
75. Lennernas H. Human in vivo regional intestinal permeability: importance for pharmaceutical drug development. *Mol Pharm* 2014;**11**:12–23.
76. Tannergren C, Bergendal A, Lennernas H, Abrahamsson B. Toward an increased understanding of the barriers to colonic drug absorption in humans: implications for early controlled release candidate assessment. *Mol Pharm* 2009;**6**:60–73.
77. Sjögren E, Dahlgren D, Roos C, Lennernäs H. Human in vivo regional intestinal permeability: quantitation using site-specific drug absorption data. *Mol Pharm* 2015;**12**(6):2026–39.
78. Lennernas H. Regional intestinal drug permeation: biopharmaceutics and drug development. *Eur J Pharm Sci* 2014;**57**:333–41.
79. Benet LZ, Broccatelli F, Oprea TI. BDDCS applied to over 900 drugs. *AAPS J* 2011;**13**(4):519–47.
80. Olivares-Morales A, Kamiyama Y, Darwich AS, Aarons L, Rostami-Hodjegan A. Analysis of the impact of controlled release formulations on oral drug absorption, gut wall metabolism and relative bioavailability of CYP3A substrates using a physiologically-based pharmacokinetic model. *Eur J Pharm Sci* 2015;**67**:32–44 (2014)
81. Tannergren C, Borde A, Boreström C, Abrahamsson B, Lindahl A. Evaluation of an in vitro faecal degradation method for early assessment of the impact of colonic degradation on colonic absorption in humans. *Eur J Pharm Sci* 2014;**57**:200–6.
82. Gayer CP, Basson MD. The effects of mechanical forces on intestinal physiology and pathology. *Cell Signal* 2009;**21**(8):1237–44.
83. Qiu Y, Cheskin H, Engh K, Poska R. Once-a-day controlled-release dosage form of divalproex sodium I: formulation design and in vitro/in vivo investigations. *J Pharm Sci* 2003;**92**(6):1166–73.
84. Viridén A, Abrahmsén-Alami S, Wittgren B, Larsson A. Release of theophylline and carbamazepine from matrix tablets—consequences of HPMC chemical heterogeneity. *Eur J Pharm Biopharm* 2011;**78**(3):470–9.
85. Lin Z, Zhou D, Hoag S, Qiu Y. Influence of drug properties and formulation on in vitro drug release and biowaiver regulation of oral extended release dosage forms. *AAPS J* 2016;**18**(2):333–45.
86. Zahirula Z, Khan I. Dissolution testing for sustained or controlled release oral dosage forms and correlation with in vivo data: challenges and opportunities. *Int J Pharm* 1996;**140**(2):131–43.
87. Kostewicz ES, Abrahamsson B, Brewster M, Brouwers J, Butler J, Carlert S, et al. In vitro models for the prediction of in vivo performance of oral dosage forms. *Eur J Pharm Sci* 2014;**57**:342–66.
88. Blanquest S, et al. A dynamic artificial gastrointestinal system for studying the behavior of orally administered drug dosage forms under various physiological conditions. *Pharm Res* 2004;**21**(4):585–9.
89. Minekus M, et al. A computer-controlled system to simulate conditions of the large intestine with peristaltic mixing, water absorption and absorption of fermentation products. *Appl Microbiol Biotechnol* 1999;**53**:108–14.
90. Barker R, Abrahamsson B, Kruusmägi M. Application and validation of an advanced gastrointestinal in vitro model for the evaluation of drug product performance in pharmaceutical development. *J Pharm Sci* 2014;**103**(11):3704–12.
91. Dickinson PA, Rmaileh RA, Ashworth L, Barker RA, Burke WM, Patterson CM, et al. An Investigation into the utility of a multi-compartmental, dynamic, system of the upper gastrointestinal tract to support formulation development and establish bioequivalence of poorly soluble drugs. *AAPS J* 2012;**14**(2):196–205.
92. Sjögren E, Abrahamsson B, Augustijns P, Becker D, Bolger MB, Brewster M, et al. In vivo methods for drug absorption—comparative physiologies, model selection, correlations with in vitro methods (IVIVC), and applications for formulation/API/excipient characterization including food effects. *Eur J Pharm Sci* 2014;**57**:99–151.
93. Humbert, et al. In vitro-in vivo correlation of a modified-release oral form of ketotifen: in vitro dissolution rate specification. *J Pharm Sci* 1994;**83**(2):131–6.
94. Qiu Y, Cheskin H, Briskin J, Engh K. Sustained-release hydrophilic matrix tablets of zileuton: formulation and in vitro/in vivo studies. *J Controlled Release* 1997;**45**:249–56.
95. Kostwicz ES, Brauns U, Becker R, Dressman JB. Forecasting the oral absorption behavior of poorly soluble weak bases using solubility and dissolution studies in biorelevant media. *Pharm Res* 2002;**19**(3):345–9.
96. Ku M, et al. Optimization of sustained-release diltiazem formulations in man by use of an in vitro/in vivo correlation. *J Pharm Pharmacol* 1998;**50**:845–50.
97. Mu X, et al. Development and evaluation of bio-dissolution systems capable of detecting the food effect on a polysaccaride-based matrix system. *J Control Release* 2003;**93**:309–18.

98. Katori N, et al. Estimation of agitation intensity in the GI tract in humans and dogs based on in vitro/in vivo correlation. *Pharm Res* 1995;**12**(2):237–43.
99. Shameem M, et al. Oral solid controlled release dosage forms: role of GI-mechanical destructive forces and colonic release in drug absorption under fasted and fed conditions in humans. *Pharm Res* 1995;**12**(7):1049–53.
100. Aoki, et al. Evaluation of the correlation between in vivo and in vitro release of phenylpropanolamine HCl from controlled release tablets. *Int J Pharm* 1992;**85**:65–73.
101. Macheras P, et al. An in vitro model for exploring CR theophylline sustained or controlled release systems: in vitro drug release profiles. *Int J Pharm* 1989;**54**:123–30.
102. Wingstrand, et al. Bioavailability from felodipine extended release tablets with different dissolution properties. *Int J Pharm* 1990;**60**:151–6.
103. Grundy JS, et al. Studies on dissolution testing of the nifedipine gastrointestinal therapeutic system. I. Description of a two-phase in vitro dissolution test. *J Control Release* 1997;**48**:1–8.
104. Pillay V, Fassihi R. Unconventioal dissolution methodologies. *J Pharm Sci* 1999;**88**:843–51.
105. Parojčić J, Vasiljević D, Ibrić S, Djurić Z. Tablet disintegration and drug dissolution in viscous media: paracetamol IR tablets. *Inter J Pharm* 2008;**355**(1-2):93–9.
106. Nicolaides E, Symillides M, Dressman JB, Reppas C. Biorelevant dissolution testing to predict the plasma profile of lipophilic drugs after oral administration. *Pharm Res* 2001;**18**(3):380–8.
107. Vertzoni M, Fotaki N, Kostewicz E, Stippler E, Vertzoni M, Leuner C, et al. Dissolution media simulating the intralumenal composition of the small intestine: physiological issues and practical aspects. *J Pharm Pharmacol* 2004;**56**(4):453–62.
108. Al-Behaisi S, Antalb I, Morovjána G, Szúnyoga J, Drabanta S, Martonb S, et al. In vitro simulation of food effect on dissolution of deramciclane film-coated tablets and correlation with in vivo data in healthy volunteers. *Eur J Pharm Sci* 2002;**15**(2):157–62.
109. Müllertz A. Biorelevant dissolution media. In: Augustijns P, Brewster ME, editors. *Solvent systems and their selection in pharmaceutics and biopharmaceutics*. New York, NY: Springer; 2007. pp. 151–77.
110. Jones HM, Parrott N, Ohlenbusch G, Lave T. Predicting pharmacokinetic food effects using biorelevant solubility media and physiologically based modelling. *Clin Pharmacokinet* 2006;**45**(12):1213–26.
111. Pedersen BL, Brøndsted H, Lennernäs H, Christensen FN, Müllertz A, Kristensen HG. Dissolution of hydrocortisone in human and simulated intestinal fluids. *Pharm Res* 2000;(17):s.183.
112. Dressman JB, Reppas C. In vitro–in vivo correlations for lipophilic, poorly water-soluble drugs. *Eur J Pharm Sci* 2000;**11**(S2):S73–80.
113. El-Arini, et al. Theophylline controlled release preparations and fatty food: an in vitro study using the rotating dialysis cell method. *Pharm Res* 1990;**7**:1134–40.
114. Morita R, et al. Development of a new dissolution test method for an oral controlled release preparation, the PVA swelling controlled release system (SCRS). *J Control Release* 2003;**90**(1):109–17.
115. Minekus M, Havenaar, R. In vitro model of an in vivo digestive tract. US patent 5,525,305; 1996.
116. Blanquet S, Zeijdner E, Beyssac E, Meunier J, Denis S, Havenaar R, et al. A dynamic artificial gastrointestinal system for studying the behavior of orally administered drug dosage forms under various physiological conditions. *Pharm Res* 2004;**21**:37–49.
117. Yoshikawa Y, Hu Z, Kimura G, Murakami M, Yoshikawa H, Takada K. A dissolution test for a pressure-controlled colon delivery capsule: rotating beads method. *J Pharm Pharmacol* 1999;**51**(9):979–89.
118. Gu CH, Rao D, Gandhi RB, Hilden J, Raghavan K. Using a novel multicompartment dissolution system to predict the effect of gastric pH on the oral absorption of weak bases with poor intrinsic solubility. *J Pharm Sci* 2005;**94**(1):199–208.
119. He X, Kadomura S, Takekuma Y, Sugawara M, Miyazaki K. An in vitro system for prediction of oral absorption of relatively water-soluble drugs and ester prodrugs. *Int J Pharm* 2003;**263**:35–44.
120. Ann Newman A, Knipp G, Zografi G. Assessing the performance of amorphous solid dispersions. *J Pharm Sci* 2012;**101**(4):1355–77.
121. Law D, Schmitt EA, Marsh K, Everitt EA, Wang W, Fort JJ, et al. Ritonavir-PEG 8000 amorphous solid dispersions: in vitro and in vivo evaluation. *J Pharm Sci* 2004;**93**:563–70.
122. Li S, He H, Parthiban LJ, Yin H, Serajuddin ATM. IV-IVC considerations in the development of immediate-release oral dosage form. *J Pharm Sci* 2006;**94**(7):1396–417.
123. Qiu Y. Case studies: application of IVIVR based on product and process understanding. AAPS/FDA/EUFEPS Workshop: AAPS Workshop on Facilitating Oral Product Development and Reducing Regulatory Burden through Novel Approaches to Assess Bioavailability/Bioequivalence. Washington, DC, 2011.
124. Qiu Y, Zhang G. Research and development aspects of oral controlled release systems. In: Wise DL, Klibanov AM, Langer R, Mikos AG, Peppas NA, Trantolo DJ, Wnek GE, Yaszeski MJ, editors. *Handbook of pharmaceutical controlled release technology*. New York, NY: Marcel Dekker, Inc; 2000. pp. 465–503.
125. Shugarts S, Benet LZ. The role of transporters in the pharmacokinetics of orally administered drugs. *Pharm Res* 2009;**26**(9):2039–54.
126. Physicians Desk Reference 2005, 59th edition. Thomson Healthcare. Stamford, CT.
127. T. R. Ju AAPS Short Course. Developing IVIVC for Oral MR Products: Drug Property, Formulation Design and Dissolution Considerations. Salt lake City, UT; 2003.
128. Pitsiu M, Sathyan G, Gupta S, Verotta D. A semiparametric deconvolution model to establish in vivo-in vitro correlation applied to OROS oxybutynin. *J Pharm Sci* 2001;**90**(6):702–12.
129. Drover DR, Angst MS, Valle M. Input characteristics and bioavailability after administration of immediate and a new extended-release formulation of hydromorphone in healthy volunteers. *Anesthesiology* 2002;**97**:827–36.
130. Wang Y, Lee L, Somma R, Thompson G, Bakhtiar R, Lee J, et al. In vitro dissolution and in vivo oral absorption of methylphenidate from a bimodal release formulation in healthy volunteers. *Biopharm Drug Dispos* 2004;**25**(2):91–8.
131. Sako K, Sawadaa T, Nakashimaa H, Yokohamaa S, Sonobe T. Influence of water soluble fillers in hydroxypropylmethylcellulose matrices on in vitro and in vivo drug release. *J Control Release* 2002;**81**(1-2):165–72.
132. Verbekena D, Neirinckb N, Van Der Meerenb P, Dewettinck K. Influence of κ-carrageenan on the thermal gelation of salt-soluble meat proteins. *Meat Sci* 2005;**70**(1):161–6.
133. Melia CD, Rajabi-Siahboomic AR, Bowtell RW. Magnetic resonance imaging of controlled release pharmaceutical dosage forms. *Pharm Sci Technol Today* 1998;**1**(1):32–9.
134. Richardsona JC, Bowtellb RW, Mäderc K, Melia CD. Pharmaceutical applications of magnetic resonance imaging (MRI). *Adv Drug Deliv Rev* 2005;**57**(8):1191–209.

135. Roshdy M, Schnaare RL, Sugita ET, Schwartz JB. The effect of controlled release tablet performance and hydrogel strength on in vitro/in vivo correlation. *Pharm Dev Technol* 2002;**7**(2):155–68.
136. Bajwa GS, Hoebler K, Sammon C, Timmins P, Melia CD. Microstructural imaging of early gel layer formation in HPMC matrices. *J Pharm Sci* 2006;**95**(10):2145–57.
137. Yang L, Johnson B, Fassihi R. Determination of continuous changes in the gel layer thickness of poly(ethylene oxide) and HPMC tablets undergoing hydration: a texture analysis study. *Pharm Res* 1998;**15**(12):1902–6.
138. Meyvisa TKL, Stubbea BG, Van Steenbergenb MJ, Henninkb WE, De Smedt SC, Demeestera J. A comparison between the use of dynamic mechanical analysis and oscillatory shear rheometry for the characterisation of hydrogels. *Inter J Pharm* 2002;**244**(1-2):163–8.
139. Grundy JS, Anderson KE, Rogers JA, Foster RT. Studies on dissolution testing of the nifedipine gastrointestinal therapeutic system. II. Improved in vitro-in vivo correlation using a two-phase dissolution test. *J Controlled Release* 1997;**48**(1):9–17.
140. Hegasy A, Ramsch K. Solid medicament formulations containing nifedipine, and processes for their preparation. United States Patent 5,264,446.
141. Katzhendler I, Azoury R, Friedman M. Crystalline properties of carbamazepine in sustained release hydrophilic matrix tablets based on hydroxypropyl methylcellulose. *J Control Release* 1998;**54**(1):69–85.
142. Li X, Zhi F, Hu Y. Investigation of excipient and processing on solid phase transformation and dissolution of ciprofloxacin. *Int J Pharm* 2007;**328**(2):177–82.
143. Qiu Y, Li X, Duan JZ. Influence of drug property and product design on in vitro-in vivo correlation of complex modified-release dosage forms. *J Pharm Sci* 2014;**103**(2):507–16.
144. Rohrs BR. Dissolution assay development for in vitro-in vivo correlations. *Am Pharm Rev* 2003;**6**(1):8–12.
145. Takano R, Kataoka M, Yamashita S. Integrating drug permeability with dissolution profile to develop IVIVC. *Biopharm Drug Dispos* 2012;**33**(7):354–65.
146. Takeuchi S, Tsume Y, Amidon GE, Amidon GL. Evaluation of a three compartment in vitro gastrointestinal simulator dissolution apparatus to predict in vivo dissolution. *J Pharm Sci* 2014;**103**(11):3416–22.
147. Guidance for Industry (draft): Dissolution Testing and Specification Criteria for Immediate-Release Solid Oral Dosage Forms Containing Biopharmaceutics Classification System Class 1 and 3 Drugs. U.S. Department of Health, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), August, 2015.
148. Roudier B, Davit BM, Beyssac E, Cardot JM. In vitro- in vivo correlation's dissolution limits setting. *Pharm Res* 2014;**31**(9):2529–38.
149. Lake OA, Olling M, Barends DM. In vitro/in vivo correlations of dissolution data of carbamazepine immediate release tablets with pharmacokinetic data obtained in healthy volunteers. *Eur J Pharm Biopharm* 1999;**48**(1):13–19.
150. Piscitelli DA, Young D. Setting dissolution specifications for modified-release dosage forms. *Adv Exp Med Biol* 1997;**423**:159–66.
151. Uppoor VRS. Regulatory perspectives on in vitro (dissolution)/in vivo (bioavailability) correlations. *J Control Release* 2001;**72**(1-3):127–32.
152. Elkoshi Z. Dissolution specifications based on release rates. *J Pharm Sci* 1999;**88**(4):434–44.
153. Hayes S, Dunne A, Smart T, Davis J. Interpretation and optimization of the dissolution specifications for a modified release product with an in vivo-in vitro correlation (IVIVC). *J Pharm Sci* 2004;**93**(3):571–81.
154. Guidance for industry, waiver of in vivo bioavailability and bioequivalence studies for immediate-release solid oral dosage forms based on a biopharmaceutics classification system. U.S. Department of Health, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), August, 2000.
155. Guidance for industry (draft): waiver of in vivo bioavailability and bioequivalence studies for immediate-release solid oral dosage forms based on a biopharmaceutics classification system. U.S. Department of Health, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), May, 2015.
156. Guideline on the investigation of bioequivalence European Medicines Agency, COMMITTEE FOR Medicinal Products For Human Use (CHMP), CPMP/EWP/QWP/1401/98 Rev. 1/Corr. London,U.K., January, 2010.
157. Biopharmaceutics Classification System (BCS) Guidance. U.S. Department of Health, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Office of Pharmaceutical Sciences. <http://www.fda.gov/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/ucm128219.htm>.
158. Biowaiver Monographs. International Pharmaceutical Federation (FIP), Sig on Regulatory Sciences, Focus Group on BCS and Biowaivers. <http://www.fip.org/bcs_monographs; http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)1520-6017/homepage/chichester.htm>.
159. Guidance for industry: SUPAC-MR: modified release solid oral dosage forms scale-up and post approval changes: chemistry, manufacturing, and controls; in vitro dissolution testing and in vivo bioequivalence documentation. U.S. Department of Health, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), September, 1997.
160. Siewert M. Perspectives of *in vitro* dissolution test in establishing *in vivo/in vitro* correlations. *Eur J Drug Metab Pharmacokinet* 1993;**18**(1):7–18.
161. Cardot JM, Beyssac E. *In vitro/in vivo* correlations: scientific implications and standardisation. *Eur J Drug Metab Pharmacokinet* 1993;**18**(1):113–20.
162. Sirisuth N, Eddington ND. The influence of first pass metabolism on the development and validation of an IVIVC for metoprolol extended release tablets. *Eur J Pharm Biopharm* 2002;**53**(3):301–9.
163. Mistry B, Leslie JL, Eddington ND. Influence of input rate on the stereospecific and nonstereospecific first pass metabolism and pharmacokinetics of metoprolol extended release formulations. *Chirality* 2002;**14**(4):297–304.
164. Qiu Y, Law D, Krill S. Use of total plasma concentration of a drug given as racemate in absorption analysis is not valid for drug exhibiting stereoselective disposition and/or absorption. Proc. 22nd Int. Sym. on Controlled Release of Bioactive Materials 280; 1995.
165. Cao GL, Locke C. Assessing whether controlled release products with differing in vitro dissolution rates have the same in vivo-in vitro relationship. In: Young D, Devane J, Butler J, editors. *In vitro-in vivo correlations*. New York, NY: Plenum Press; 1997. pp. 173–80.
166. Kortejärvi H, Malkki J, Marvola M, Urtti A, Yliperttula M, Pajunen P. Level A in vitro-in vivo correlation (IVIVC) model with Bayesian approach to formulation series. *J Pharm Sci* 2007;**95**(7):1595–605.
167. Grundy JS, Foster RT. The nifedipine gastrointestinal therapeutic system (GITS). Evaluation of pharmaceutical, pharmacokinetic and pharmacological properties. *Clin Pharmacokinet* 1996;**30**(1):28–51.

168. Schug BS, Brendel E, Chantraine E, Wolf D, Martin W, Schall R, et al. The effect of food on the pharmacokinetics of nifedipine in two slow release formulations: pronounced lag-time after a high fat breakfast. *Br J Clin Pharmacol* 2002;**53**(6):582–8.
169. Qiu Y. Strategy and approach in developing in vitro-in vivo correlations. Proc. 3rd CPA/AAPS Joint Sym. on: "Modern Theory and Practice of Pharmaceutical Sciences in Research and Development of Oral Dosage Form". Shanghai; 2005. pp. 261–270.
170. Qiu Y. Mitigating patient risks using dissolution testing in manufacturing extended release products. 2nd FDA/PQRI Conference on Advancing Product Quality, Rockville, MD; 2015.
171. Wojcik RC. Setting dissolution specifications. Dissolution Tech. 12-20. AUGUST 1997.
172. Banholzer M, Gong,Y, Apfelbaum A, Xiong H and Qiu Y. Setting dissolution specifications based on in vitro/in vivo correlation (IVIVC) and Monte Carlo simulation. AAPS Annual Meeting and Exposition, Orlando, FL; 2015.
173. Dong X, Tsong Y, Shen M. Statistical considerations in setting product specifications. *J Biopharm Stat* 2015;**25**(2):280–94.
174. Monte Carlo Analysis. Module KQA: knowledge quality assessment. The Forest of Broceliande; 2016. <https://proxy.eplanete.net/galleries/broceliande7/monte-carlo-analysis>.
175. Monte Carlo method. Wikipedia.
176. Duan J, et al. The application of IVIVC in setting the clinical relevant manufacturing specifications. AAPS Annual Meeting. San Antonio, TX; 2013.
177. Gualtieri, et al. Clinical studies of methylphenidate serum levels in children and adults. *J Am Academy Child Psychiatry* 1982;**21**:19–26.

PART III

DESIGN, DEVELOPMENT AND SCALE-UP OF FORMULATION AND MANUFACTURING PROCESS

This page intentionally left blank

CHAPTER

17

Oral Formulations for Preclinical Studies: Principle, Design, and Development Considerations

Y. Gao[1], *C. Gesenberg*[2] *and W. Zheng*[3]

[1]AbbVie, Inc., North Chicago, IL, United States [2]Bristol-Myers Squibb, New York City, NY, United States
[3]AstraZeneca Pharmaceuticals LP, Waltham, MA, United States

17.1 INTRODUCTION

Preclinical studies are designed to verify therapeutic targets and assess the efficacy and safety of new drug candidates in relevant in vitro and in vivo models. The goals of preclinical studies are to select the right drug candidates to test in humans and understand the translation from preclinical results to clinical outcomes.

Preclinical in vivo studies include pharmacokinetics (PK), efficacy, and safety studies. Early efficacy studies, including pharmacodynamics (PD), dose scheduling, and combination studies, are usually conducted in mice and rats. Safety evaluations, including maximum tolerated dose (MTD), dose range finding, and good laboratory practice (GLP) studies, are conducted in rodent and nonrodent species. Furthermore, telemetry studies and other in vivo assessments may be required to mitigate development risks at the preclinical stage.

Different routes of administration can be explored at the preclinical stage for hypothesis testing. Intravenous (IV) bolus, IV infusion, oral, subcutaneous and intraperitoneal administrations are commonly used at the early drug discovery stage to understand exposure and signal relationships. However, the same administration route as the intended route for the clinical and commercial drug products is typically required for efficacy and safety evaluations of drug candidates. Oral and IV infusion are the most common administration routes for preclinical safety evaluations.

A desired preclinical formulation should provide sufficient and robust exposure, have fit-for-purpose stability, and not cause adverse effects or interference of test signals in animal models. Achieving the appropriate exposure during the preclinical stage is important, as it allows evaluations of target engagement, PK/PD and efficacy relationships, and early safety liabilities. Suitable formulations are critical to ensure the achievement of the appropriate exposure for in vivo assessments. Rational preclinical formulation development can also facilitate candidate selection and guide clinical formulation development.

The common challenges of achieving optimal exposure and formulation development at the preclinical stage include: (1) early leads often have suboptimal physicochemical (eg, poor solubility, purity, stability and solid-state properties, batch-to-batch variation) and PK properties (eg, low permeability, high clearance); (2) lack of PK/PD information and a clear understanding of the optimal exposure or coverage; (3) limited compound availability and short timelines; and (4) lack of in vitro and in vivo relationship for formulation selection.

Preclinical formulation efforts are often focused on three areas: (1) enhance solubility and exposure for poorly soluble compounds, (2) reduce C_{max} and/or prolong exposure to improve the therapeutic index, and (3) increase exposure at the site of action by local delivery and targeting.

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00017-0

This chapter is centered on improving oral exposure through rational formulation design. There are six sections in this chapter:

1. Considerations in Designing Formulations for Preclinical Species
   Review formulation, study requirements, and considerations for safety evaluation.
2. Use of Active Pharmaceutical Ingredient (API) Properties to Guide Formulation Design
   Discuss key physicochemical properties and their impact on formulation design and performance.
3. Preclinical Formulation Development
   Detail preclinical formulation approaches with a focus on enabling technologies for poorly water-soluble compounds.
4. In Vitro Dissolution Method to Evaluate Formulations
   Describe current advancements in understanding of dissolution behaviors of enabling formulation and rational design of in vitro dissolution methods.
5. Preclinical Formulation Selection
   Discuss pros and cons of formulation approaches and rational formulation selection.
6. Case Study

Illustrate the formulation development process using a compound from real world.

17.2 CONSIDERATIONS IN DESIGNING FORMULATIONS FOR PRECLINICAL SPECIES

17.2.1 Type and requirements of nonclinical safety assessment studies

A nonclinical safety assessment usually begins in the drug discovery phase as part of the screening process that leads to the selection of a candidate for clinical development, and it reaches all the way into late-stage clinical development of a potential new drug. The main purpose of non-GLP studies during drug discovery is the identification of development-limiting toxicity. Studies are usually performed for the last 1–3 drug discovery lead compounds in a rodent and nonrodent species and aid in the selection of the final compound chosen to progress into development. Single-dose studies are typically performed at three-dose levels to gain information on the toxicokinetic profile and drug tolerability and to ultimately select a dose range for repeat-dose toxicology studies. Typical doses for these early single-dose studies are in the range of 30–300 mg/kg. Prior to the final selection of a drug candidate for development, a repeat-dose range-finding study is conducted for typically 7–14 days in a rodent (rat or mouse) and nonrodent (dog or monkey) species, with the route of administration usually being driven by the final intended route of administration in humans.

During the development phase, single and repeat-dose studies are designed to assess various toxicology areas such as carcinogenicity, immunotoxicity, phototoxicity, reproductive and developmental toxicity, genotoxicity, local tolerance, and abuse liability. The typical endpoints for repeat-dose toxicology studies are toxicokinetics, clinical signs, clinical pathology (hematology, serum chemistry, coagulation, urinalysis), gross pathology, histopathology, and safety pharmacology (cardiovascular, central nervous system, respiratory). Core safety pharmacology and genetic toxicology studies need to be conducted prior to a new drug candidate entering clinical studies in humans for the first time. Timing for reproductive studies depends on the human population to be exposed. Prior to the compound entering the first clinical studies, an investigational new drug-enabling GLP study is required in a nonrodent and a rodent species, which usually lasts for 4 weeks, and is needed to set clinical starting doses and dose-escalation paradigms. Safety studies in later development stages in general provide guidance to clinical investigators and maintain subject and patient safety in clinical studies and following registration. A stepwise process is utilized in the planning of clinical trials to assess the potential toxicity in advance of use in humans. For short clinical studies lasting for less than 6 months, the same duration of a preclinical study in a rodent and a nonrodent species is required. For long-term clinical studies lasting for more than 6 months, a 6-month rodent and a 9-month nonrodent study are performed. Carcinogenicity studies are usually conducted in mice and rats and last for 2 years, with tumorigenicity as the main endpoint. However, for products with a maximal clinical dosage duration of less than 6 months, or those intended for life-threatening oncology indications, carcinogenicity studies are generally not required.[1]

17.2.2 Complexities caused by high exposure requirement but minimal adverse effect

The timing of a nonclinical study (discovery or development), dose regimen (dose, frequency, and duration), test species, and route of administration all have a significant impact on the formulation development strategy. The ultimate goal is to maximize the exposure in the test species to identify the safety risks associated with the test compound. The desired target plasma concentrations of nonclinical toxicology studies

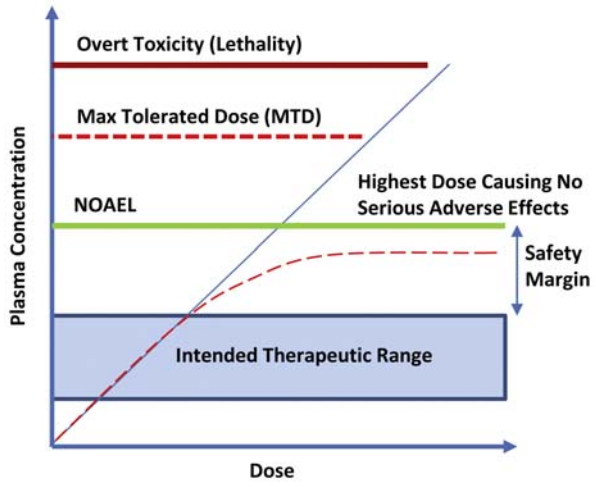


FIGURE 17.1 Interplay between dose and exposure level requirement in various studies. *MTD*, maximum tolerated dose; *NOAEL*, no observed adverse effect level.

usually far exceed the intended therapeutic range for a given drug, oftentimes 100-fold or more, depending on the therapeutic area for which the drug is developed. Especially in rodent species, extremely high doses are often explored to establish a no observed adverse effect level, the highest dose causing no serious adverse effect; an MTD; and an overt toxicity level (lethality). In order to enable the extremely high plasma concentrations desired in nonclinical studies, the formulator is challenged with the identification of formulation options that solubilize the drug substance and keep it from precipitating after administration to achieve a close-to-dose proportional increase in exposure over a very wide dose range. However, due to the limited solubility of most test compounds, a less-than-dose proportional increase in exposure and ultimately a plateau beyond which exposure does not increase further with increasing dose are often observed at the higher end of the dose range, as shown in Fig. 17.1.

17.2.3 Complexities in dosing preclinical species

At the same time, a formulator needs to be aware of the physiological differences between nonclinical test species, some of which are summarized in Table 17.1. Especially for oral administration, these physiological differences can have a profound effect on the performance of a formulation, since the oral absorption is, in many cases, limited by the test compound's solubility or dissolution rate, and, therefore, factors such as gastric residence time and gastrointestinal (GI) residence time

TABLE 17.1 Physiological Differences Between Nonclinical Test Species[2–12]

Parameters	Humans	Rats	Dogs	Monkeys
Fasted gastric pH	0.95–3.2	1.5–3.0	3.5–4.5	2.8–4.0
Fed gastric pH	4.3–5.4	3.8–5.0	3.5–5.5	
Basal gastric output (mEq/h)	2.0–5.0	0.8657 (mL/4 h/ 100 g)[13]	0.1	Similar to humans unless stressed
Peak output (fed state), (mEq/h)	18–23		39	Similar to humans unless stressed
Duodenum pH	5.0–7.0	6.5–7.0	5.5–7.5	5.6–6.0
Stomach capacity (L)	1–1.6		1.0	0.1
Gastric emptying time (min)	16–140	10–30	100	
Sm. intestine transit time (min)	240	90	60–120	
Total GI residence time (min)	2400	1200	770	1200

can greatly impact the exposure achieved from a given formulation in different species. In addition, since many test compounds are ionizable, the differences in the gastric and intestinal pH of the test species could cause significant differences in the ionization state and solubility of the test compound, leading to differences in the tendency or extend of precipitation from a solubilized formulation once it enters the gastric milieu after oral administration.

Differences in the stomach capacity of test species result in species-specific restrictions to the maximum volume that can be administered orally per dose. The administration of excessive dose volumes may produce pain, excitement, altered physiological parameters (eg, serum electrolyte imbalance, increased blood pressure, and increased respiration rate) and can cause abnormal test compound absorption. The ideal and maximum dose volume limits are usually defined by animal care and use committees in each organization that conducts animal studies. While these limits are typically aligned based on surveys amongst pharmaceutical companies, biotech companies, and contract research organizations and based on literature,[14] there might be differences in the limits used from organization to organization. Table 17.2 lists commonly used maximum dose volumes for different species that the authors have experience with, but these should not be interpreted as absolute guidelines.

TABLE 17.2 Dose Volume Guidance

Species	Average Weight (kg)	Dose Volume (mL/kg)	Average Max. Volume per Dose (mL)	Solution/ Suspension Conc. Required Using Example Dose of 20 mg/kg (mg/mL)
Mouse	0.025	10	0.25	2
Rat	0.25	10	2.5	2
Rabbit	4	10	40	2
Dog	10	5	50	4
Macaque (Cyno)	5	5	25	4
Marmoset	0.4	10	4	2

These dose volume restrictions have a significant impact on the solution or suspension concentration that is required to administer a defined dose. Therefore, the target concentration that a formulator has to aim for to administer a given dose can vary depending on what test species the formulation is developed for, as outlined in Table 17.2.

17.2.4 Complexities due to use-limit of excipients

An additional challenge for the formulation selection for nonclinical safety studies is the fact that the majority of newly selected drug candidates have poor aqueous solubility and low bioavailability (BA) from suspension formulations, and oftentimes the formulator needs to employ a combination of multiple cosolvents, surfactants, or other excipients to increase the solubility and the precipitation resistance of the test compound to achieve the desired in vivo exposures. However, it is important to consider the potential for undesirable effects of the selected excipients and to stay within the ranges that do not lead to toxicity findings caused by the vehicle itself, which could interfere with the intended readout of the safety study.[1,15,16] Excipients employed in formulations should ideally maximize exposure while having minimal impact on the pharmacology and toxicology readouts. The final dosing formulation will be selected based on the chemical and physical stability, the solubility, the required concentration (mg/mL), wettability, crystallinity, and the acceptability of the excipients in animals. Since many excipients have physiological effects (eg, GI permeability, transit time, mobility, efflux, and metabolism), their impact on the PK, efficacy, and toxicity should be taken into consideration. At times, the side effects observed after the administration of the vehicle alone may appear manageable, but these effects could exacerbate drug-related side effects, thereby complicating the interpretation of the study results.

The acceptable level of a specific excipient also depends on the dose frequency and duration of the study. For single-dose studies or repeat-dose studies that only span several days, a well-understood, low-level biological or toxic effect of the vehicle might be acceptable, allowing the formulator to use higher concentrations and doses of certain excipients. However, for long-term repeat-dose studies, such effects may not be tolerated, and the formulation has to be adjusted to lower levels of the excipient that causes the effect.

When selecting a vehicle, both the concentration and the total dose of the excipients should be taken into consideration. The sensitivity towards a specific excipient can be species-dependent. A common example of a vehicle-related effect is the laxative effect of the cosolvent polyethylene glycol (PEG) 400 or surfactants such as poloxamers, causing loose stool and emesis in dogs and primates, which can lead to variable and reduced exposure.[17–20] Rats and mice, on the other hand, usually show a much higher tolerance for such excipients and can, therefore, be safely dosed with higher levels. Surfactants such as poloxamers, Cremophor EL or vitamin E D-alpha tocopheryl polyethylene glycol 1000 succinate (TPGS) at high doses may inhibit the efflux transporters such as P-gp, and, therefore, have the potential to impact the PK profile of certain test compounds that are P-gp substrates. In addition, they may alter the extent of gut wall metabolism and change the exposure of metabolites.[17,21] The acceptability and potential side effects of a given excipient can also be related to the intended study duration. Corn oil, for example, is well suited for short-term studies, but when administered chronically in rats, it can cause altered body weight, survival rate, and tumor incidence.[1] Vehicle-related toxicities or changes in PK or metabolite profile of the test compound could complicate the interpretation of the study results, especially when there is no previous internal experience or published data. As with dose-volume restrictions mentioned earlier, many organizations have established their own internal guidance on the acceptable limits for each of the commonly used excipients depending on test species (and in some cases different strains of the same species), route of administration, and dose regiment (single or repeat-dose study).

Table 17.3 provides examples of such limits for typically used levels of certain excipients and commonly observed adverse effects associated with these excipients that are based on literature[1,15–17,22] and experience from the authors' organizations.

TABLE 17.3 Examples of Excipients Use Limits

		Single Dose								
		Mice		Rats		Dogs		Monkeys		
		Oral % w/w (g/kg)	IV % w/w (g/kg)	Oral % w/w (g/kg)	IV % w/w (g/kg)	Oral % w/w (g/kg)	IV % w/w (g/kg)	Oral % w/w (g/kg)	IV % w/w (g/kg)	Comments
Cosolvents	PEG 400 (Polyethylene glycol 400)	100 (10)	50 (1.5)	100 (20)	50 (1.5)	80 (5)	30 (1)	25 (5)	30 (1)	Loose stool, emesis, renal toxicity
	PG (Propylene glycol)	80 (2)	80 (1)	80 (6)	80 (1)	80 (5)	80 (5)	50 (5)	50 (5)	Loose stool, emesis
	DMSO (dimethyl sulfoxide)	50 (0.5)	100 (0.1)	50 (0.5)	100 (0.1)	50 (0.5)	100 (0.05)	50 (0.5)	NR	
	Ethanol	20 (1)	20 (0.5)	20 (1)	20 (0.5)	50 (2)	20 (0.5)	25 (2)	20 (0.1)	Peritoneal adhesions, liver and spleen peritonitis
	DMA (*N*,*N*-dimethylacetamide)	10 (1)		10 (1)		10 (0.5)	2(0.05)	10 (0.5)	2(0.1)	
Surfactants	Tween 80 (polysorbate 80 / polyoxyethylene sorbitan monooleate)	10 (2.5)	2.5 (0.8)	10 (2.5)	2.5 (0.3)	10 (1)	NR	10 (1)	NR	Anaphylaxis in dogs when dosed IV
	Pluronic F68 (Polyoxyethylene-polyoxypropylene block copolymer)	20 (2)	20 (1)	20 (1.5)	20 (1)	15 (1.5)	15 (0.5)	15 (1)	15 (0.5)	
	Cremophor EL (polyoxyl castor oil)	10 (3)	10 (0.3)	10 (3)	10 (0.3)	10 (3)	10 (0.1)	10 (0.5)	NR	Loose stool, anaphylaxis in dogs when dosed IV
	Labrasol (caprylocaproyl macrogol-8-glycerides)	50	30	50	30	40 (1)		70 (2)		
	Vitamin E TPGS 1000 (D-α-tocopheryl polyethylene glycol 1000 succinate)			25 (2.5)		25 (1)				Causes hemolysis if administered IV
Cyclodextrins	HP-β-CD (hydroxypropyl-β-cyclodextrin)	50 (5)	50 (2.5)	50 (5)	50 (2.5)	50 (5)	50 (2.5)	50 (2.5)	50 (0.5)	Loose stool
	SB-β-CD (sulfobutylether-β-cyclodextrin)	30 (5)	30 (4)	30 (5)	30 (4)	30 (1.5)	20 (3)	30 (1)	30 (3)	Loose stool, renal tubule vacuolation and foamy macrophages at large doses
Lipids	Miglyol 812 (mid-chain triglyceride of caprylic/caproic acid)	60 (1.5)	30	60 (1.5)	30			15 (1.5)		Loose stool
	Labrafil 1944CS (polyoxyethylated oleic glycerides)	4		4		4		4		Loose stool
	Capmul MCM (medium chain mono- and diglycerides)	50	40	50 (1)	40	2.5				Loose stool

17.3 USE OF API PROPERTIES TO GUIDE FORMULATION DESIGN

Understanding of API physicochemical properties is essential for developability assessment and formulation design. Oral BA can be impacted by many factors, including solubility, dissolution rate, stability, permeability, and metabolism. It is critical to understand the BA limiting factors and if exposure can be improved through formulation design. For example, if the oral exposure is limited by solubility/dissolution or stability, formulation can usually improve the exposure. However, if the oral exposure is limited by permeability or metabolism, then it is less likely to be improved through formulation.

In this section, we will discuss the impact of key API properties on oral absorption and formulation design.

17.3.1 Solubility and bioavailability

Solubility is one of the key attributes of a drug candidate, as it can significantly impact in vitro profiling, in vivo exposure, and formulation development and performance. Low solubility is one of the major challenges for drug discovery and development. Up to 75% of drug candidates in development are classified as low solubility.[23] It is believed that high-throughput screen (HTS)-based chemistry approaches and exploration of more challenging targets are leading to more hydrophobic and less soluble drug candidates.[23–25]

17.3.1.1 Factors that impact solubility

Aqueous solubility of a given molecule is the interplay of multiple factors, including solid-state properties (eg, crystal packing, lattice energy), ionization (pH, pK_a), and solute/solvate interactions. It is important to understand what limits solubility to guide the formulation effort.

17.3.1.2 Solid-state properties

The equilibrium solubility of a pharmaceutical solid is not only dependent on solvent and solute interactions but also on the intermolecular interactions in the solid state. The compound has to overcome crystal packing to dissolve in the solvent. Therefore, solid-state properties can impact API stability and solubility. Higher apparent solubility is typically observed from solids with a higher energy state (eg, metastable or amorphous solids). The most thermodynamically stable form has the lowest energy stage and the lowest solubility. The melting point and heat of fusion can be reasonable indicators for lattice energy at early development. Disruption of the tight crystal packing through the molecule design can be a viable approach to increase the solubility at the drug discovery stage. Recently, the solid-state perturbation tool has been adopted to predict solubility and suggest chemistry modifications to improve solubility based on the interaction energy calculated from a single crystal structure.[26] The rationale of this approach is to improve solubility by altering the crystal packing energy via designing in weaker intermolecular interactions without impacting the other key properties such as potency, selectivity, and lipophilicity.

17.3.1.3 pH and pK_a

The ionization state of a molecule can significantly impact its solubility. The solubility of an ionizable compound is strongly dependent on the pH of the media and the pK_a of the molecule. The solubility of an ionizable compound can be described by the pH-solubility relationship derived from the Henderson–Hasselbalch equation, where the solubility is a function of intrinsic solubility (S_0), pH of the medium, and pK_a of the molecule. Detailed solubility equations are listed in chapter "Solubility of Pharmaceutical Solids" (Eq. 1.42 for acidic compound and Eq. 1.45 for basic compound). Fig. 17.2 shows the solubility of acidic compounds as a function of pH, intrinsic solubility, and pK_a. It highlights the importance of selecting a compound with a higher intrinsic solubility. It also indicates the benefit of having a stronger pK_a in chemical structure for solubilization purpose.

However, it is well known that the ionization state not only impacts solubility but also affects permeability,

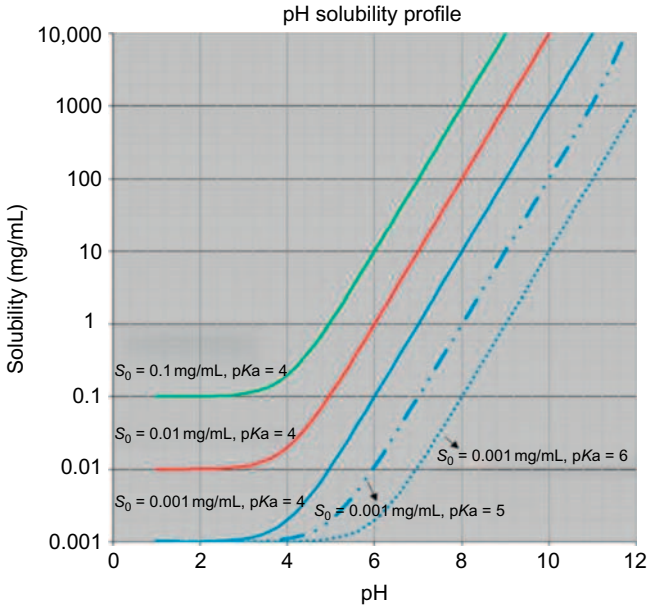


FIGURE 17.2 Illustration of how changes in intrinsic solubility S_0 and pK_a affect the pH-solubility profile of an acidic compound.

distribution, and other PK properties. The ionization state can also impact binding to transporters, enzymes, complexation (eg, with cyclodextrins),[27] and partition into organic solvents. Therefore, many properties need to be considered and balanced when it comes to the design or the selection of an optimal pK_a value for a drug candidate.

17.3.1.4 Lipophilicity

Lipophilicity plays a significant role in drug discovery and compound design. The lipophilicity of an organic compound can be described by a partition coefficient, log*P*, which can be defined as the ratio of the concentration of the unionized compound at equilibrium between organic and aqueous phases. For compounds with ionizable groups, the distribution of species is impacted by pH. The ionization state of a molecule can impact its lipophilicity. This leads to the definition of the distribution coefficient (log*D*) of a compound, which takes into account the dissociation of weak acids and bases. Generally speaking, more lipophilic compounds are less soluble in aqueous media. However, lipophilic compounds may have good solubility in oils and lipids; therefore, they can be good candidates for lipid-based formulations.

Lipophilicity not only impacts solubility but also influences permeability; potency; selectivity; absorption, distribution, metabolism, and excretion (ADME) properties; and toxicity. High lipophilicity (log$P > 5$) often contributes to high metabolic turnover, low solubility, and poor oral absorption. In addition, highly lipophilic compounds tend to bind to hydrophobic targets other than the desired target, and, therefore, there is an increased risk of promiscuity and toxicity.[28,29] Low lipophilicity can also negatively impact permeability and potency and thus results in low BA and efficacy. Leeson and St. Gallay conducted an analysis on patented compounds between 2000 and 2010 from 18 pharmaceutical companies, including Abbott, AstraZeneca, Bristol-Myers Squibb, Lilly, and Novartis. They found that the mean log*P* of these patented compounds range from 3.5 to 4.5. In addition, across the 18 companies, there was a trend of reduction in cLog*P* over time that was small but statistically significant.[30] It is generally considered that compounds with a log*P* greater than 1 or less than 4 are more likely to have optimal physicochemical and ADME properties for oral drugs.[31–33]

17.3.2 Solubility prediction and screen

17.3.2.1 Solubility prediction

Solubility prediction remains an area of interest for many researchers. Various prediction models have been reported in the literature. The models for intrinsic solubility prediction can be divided into several categories[34]: (1) fragment or group contribution-based models,[35–37] (2) log*P* and melting-based models,[38,39] and (3) statistically derived models based on 2-D or 3-D chemical descriptors.[40–42] However, the accurate prediction of solubility remains challenging. The majority of organizations still relies on experimental solubility results for compound prioritization and decision making.

17.3.2.2 Solubility screen and measurement methods

High-throughput (HT) solubility screening methods have been widely used to profile solubility during the early drug discovery stage. A large number of compounds to be profiled in various assays also requires standardized procedures in compound handling and distribution. In most cases, these are performed in 96-well or 384-well microtiter plates. For early HT solubility assays, the compounds are often dispensed by two methods. One method is to introduce the compounds as dimethyl sulfoxide (DMSO) stock solutions into aqueous media. The alternative method is to dispense the compounds as DMSO solutions, and then aqueous media are added after DMSO is removed from the plate. The final pH and solid form are typically not measured or characterized in either case. The HT solubility is determined by measuring the concentration of the saturated solutions using UV-Vis spectroscopy or liquid chromatography (UV or LC) after the solids are removed by filtration/centrifugation or by detecting precipitation formation using UV or nephelometric turbidity detector. These HT assays tend to overestimate the solubility due to the remaining organic solvent and/or formation of amorphous material during precipitation from DMSO stock solution. However, it is a very useful tool to rank the order of the compounds and flag the solubility risks at the early stage with only a small amount of compound required. In addition, the early solubility information also helps to interpret results from other in vitro assays.

During the lead optimization phase with a short list of compounds, it is recommended to conduct an equilibrium solubility study using crystalline solids with final pH and physical form measurements. This is because the pH and the physical form can significantly impact the solubility, as we previously discussed. In addition, the physical form can change during the course of solubility experiments. For examples, a free form can form a salt or a hydrate, or an amorphous material or metastable polymorph can convert to a more stable crystalline form. Therefore, it is important to understand the physical form of the initial and final solids. The most commonly used techniques to characterize the physical form of the

residual solids from solubility studies are polarized light microscopy (PLM) and powder X-ray diffraction (PXRD) (refer to chapter: Solid-State Characterization and Techniques, for techniques used for solid-state characterization).

During the drug development stage, the "shake-flask" method is typically considered as the standard method for determining the thermodynamic solubility. In the shake-flask method, a compound with known physical form (typically with most stable crystalline form) is added into a flask/vial with an excess amount to the medium, and the resulting suspension is agitated at a predetermined temperature and speed. Samples are collected at the predefined time point(s) typically at least 1–2 days. The residual solids are removed by filtration, the solubility of the compound is quantified by high-performance liquid chromatography (HPLC) or ultra-performance liquid chromatography (UPLC), the final pH is measured and recorded, and the physical form of the residual solids is characterized.

A detailed review of the solubility determination method is out of the scope of this chapter. There is a vast amount of information on this topic in literature for interested readers.[43,44]

17.3.2.3 Solubility screen in vehicles

The pH solubility measurements are typically performed during the early stage of development. Some organizations also incorporate solubility in biorelevant media in the early-stage HT screens to understand the developability risk as oral candidates. Solubility in different organic solvents is sometimes conducted to facilitate polymorph, salt selection, and process optimization. For poorly soluble compounds, additional solubility screening is typically performed with different solubilizing vehicles to guide the formulation design and developability risk assessment. Solubilizing vehicles include cosolvents, surfactants, complexation agents, and oils/lipids. The common screening excipients are listed in Table 17.4. In addition, different counterions also impact solubility due to different K_{sp} values. In situ salt screening with different counterions can provide useful information on solubility enhancement and formulation development. The theory and practical aspects of in situ salt formation to improve solubility have been discussed in the literature.[45,46]

With more technology advances such as solid dispensing, analytical and solid-state tools, there is an increasing trend to use prediction tools at early stage and incorporate HT capability in later-stage solubility measurement to increase efficiency and reduce the cost. At the same time, there is also an increasing desire to build in more focused screens with a more rational selection of excipients and conditions based on physicochemical properties of the compound.

TABLE 17.4 Typical Excipients and Vehicles Screened for Solubility During Early Stage of Development

Type of Vehicles	Examples
pH buffers	pH 1, pH 4, pH 6.5, pH 7.4, pH 9
Biorelevant media	SGF, FaSSIF, FeSSIF
Cosolvents	Polyethylene 400 (PEG 400), propylene glycol (PG), ethanol and dimethylacetamide (DMA)
Surfactants	Polysorbate 80 (Tween 80), Polyoxyl 15 Hydroxystearate (Solutol® HS 15/Kolliphor® HS 15), Pluronic® F-68, Polyoxyl-35 castor oil (Cremophor® EL/Kolliphor® EL), D-α-tocopheryl polyethylene glycol 1000 succinate (Vitamin E TPGS), caprylocaproyl polyoxyl-8 glycerides (Labrasol), sodium lauryl sulfate (SLS)
Complexation	Hydroxypropyl-β-cyclodextrin (HP-β-CD), sulfobutylether-β-cyclodextrin (SBE-β-CD/Captisol®)
Oils/Lipids	Medium-chain triglyceride (eg, Miglyol® 812), medium-chain monoglycerides & diglycerides (eg, Capmul® MCM), phospholipids (eg, Phosal® 53 MCT)
Counterions (for ionizable compounds)	Acidic compounds: NaOH, KOH, meglumine, lysine, arginine, etc. Basic compounds: HCl, methanesulfonic acid, phosphoric acid, etc.

17.3.3 Formulation design with solubility information

If exposure is limited by solubility or dissolution rate, it is critical to understand what factors limit solubility. For ionic compounds, pH adjustment and salt formulation are the most commonly used approaches to enhance solubility and dissolution. For compounds with a high melting point and tight crystal packing, a number of approaches, including cocrystal, salt, and amorphous formulations, can be explored to disrupt the crystal packing. For compounds with a strong tendency to precipitate from a solution or crystallize from an amorphous form, exicipients such as polymers or surfactants may be required to inhibit or delay precipitation or crystallization by reducing API–API interactions, molecule mobility, or supersaturation ratio. For highly lipophilic compounds, a lipid or oil-based formulation such as an emulsion or self-emulsified drug delivery systems (SEDDS) may be considered to improve solubility. In some cases, lipid or oil-based formulations may enhance permeability and promote lymphatic absorption. These solubility/dissolution enhancing approaches will be discussed in more details in the enabling technology section.

17.3.4 Stability

The physical and chemical stability of a compound is very important for developability assessment and

formulation design. The primary goal of the stability assessment at the preclinical stage is to (1) elucidate the development risk with respect to temperature, pH, light, humidity, and oxygen; (2) guide formulation design; and (3) define the storage condition and shelf life.

Liquid formulations are the most commonly used formulations for preclinical studies. The major routes of degradation are through hydrolysis, oxidation, or photochemical means. In addition, the presence of excipients such as cosolvents, complexing agents, and surfactants can also impact formulation stability. Understanding of stability and degradation mechanisms is important for selecting the dosage form, optimizing formulation variables, and defining storage conditions. During the candidate selection phase, accelerated stability studies of drug substances and formulations are typically performed in selected conditions.

17.3.4.1 Implication to formulation design

In order to mitigate the chemical stability risk via formulation design, it would be ideal to understand the rate, mechanisms, and pathways of degradation, although it may not be feasible at the early stage of drug discovery and development. Hydrolysis is affected by pH, temperature, buffer species, and ionic strength. The reactions are typically acid or base catalyzed; therefore, pH is one of the key factors influencing the stability of a compound. It is common to adjust the pH of a formulation to maximize or balance stability and solubility. Oxidation is another common degradation encountered. Autoxidation is the most common oxidative route, and its pathway can be described in three stages: initiation, propagation, and termination. In addition, hydroperoxides, common impurities in pharmaceutical excipients such as Tween 80, PEGs, and PVP, can catalyze oxidation.

Certain excipients (eg, antioxidants or complexing agents) can be used to improve chemical stability. For example, cyclodextrins have been reported to improve the photo and hydrolysis stability due to protection of functional groups that are sensitive to the degradation.[47,48] In addition to formulation pH and excipients, concentration range and storage conditions can be optimized to maximize stability. If the compound is extremely unstable in a solution, alternative formulations need to be considered. Suspensions, oil/lipid, or solid formulations may have a better chemical stability than solution formulations.

Physical instability such as precipitation of a solution, dissociation of a salt, or physical form conversion in a suspension or solid dosage forms can also negatively impact the formulation quality and in vivo exposure. Formulation approaches, composition (including concentration, pH, and excipients), and storage conditions can be explored to achieve the required stability for preclinical studies. For example, surfactant and polymeric excipients can inhibit or delay the physical form conversion in suspensions.

17.3.5 Evolution of solid forms and batch-to-batch variation

At an early stage, the synthesis route and the crystallization process are not fully developed or optimized. Batch-to-batch variation in the physical form and impurity profiles are common with early drug discovery batches. The physical form is often not optimized, and early batches tend to be amorphous or metastable forms. The impurity profiles can also change from batch to batch. The physical form can impact solubility, stability, and, therefore, BA of the drug candidate. For example, a suspension or a solution developed with a metastable form from an early batch can achieve good exposure. However, when a more stable crystalline form is produced with later batches, the suspension formulation may not provide sufficient exposure due to decreased solubility of the new, more stable crystalline form, or the compound may not be formulated in the same solution formulation developed with the early batch. The impurity profile can also impact the formulatability and chemical stability.

17.4 FORMULATIONS FOR BCS CLASS I/III COMPOUNDS

One of the biggest challenges for toxicology dosing formulation development is the high dose requirement to assess the toxicity of test compounds. Paired with species-specific restrictions on dose volume, these dose requirements often result in very high target concentrations for formulations. The biopharmaceutical classification system (BCS), introduced in 1995 by Amidon[49] and incorporated in 2000 by the Food and Drug Administration as a guidance for industry,[50] divides drug substances into four classes, depending on their solubility and permeability (Fig. 17.3). The distinction between high and low solubility compounds depends on the highest administered clinical or commercial dose of a given compound. If this dose is soluble in $\leq$250 mL of an aqueous media across the entire physiological pH range (pH 1 to 7.5), the compound is considered high solubility and is assigned to BCS Class I or III. If the highest dose is only soluble in $>$250 mL, the compound is assigned to BCS class II or IV.

Since doses required for a nonclinical safety assessment often exceed clinical/commercial doses by far, the BCS does not strictly apply to the formulation

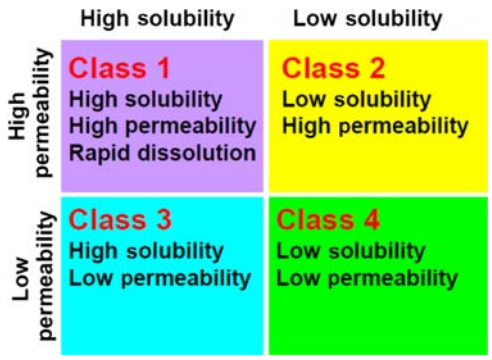


FIGURE 17.3 Biopharmaceutics classification system. *Source: Adapted based on Amidon GL, Lennernaes H, Shah VP, Crison JR. A theoretical basis for a biopharmaceutic drug classification: the correlation of in vitro drug product dissolution and in vivo bioavailability. Pharm Res 1995;12:413–20.*

development for nonclinical safety studies, since even BCS class I and III compounds do not necessarily dissolve completely in 250 mL at the highest doses administered during these studies. However, due to their high aqueous solubility, BCS class I and III compounds still represent the best chance to be formulated as simple solutions or suspensions without the need for more complicated formulation strategies involving solubility-enhancing excipients.

17.4.1 Aqueous solution formulations

The approach that oftentimes allows for the easiest formulation preparation and handling is a solution. Amongst solution formulations, purely or mostly aqueous systems are most preferred, since they eliminate the potential risk of an excipient-related toxic effect. Even if the solubility of the test compound is not sufficiently high across the entire physiological pH range, solution formulations at high concentrations might still be achievable in the case of ionizable drug substances. In addition to the reduced risk of excipient-related toxic effects, solution formulations have the advantage that solid-state properties of the drug substance, such as crystal form or particle size, are of no concern assuming the most stable form is used in producing the solution formulation. In most cases, they can be easily prepared by dissolving the drug substance at the right concentration with pH adjustment if necessary. Stirring or sonicating of the formulation for a period of time might be required to achieve full dissolution of the test compound. Stability studies that are required to support use-time recommendations for GLP studies need only be focused on the chemical stability of the test compound and potential light or temperature sensitivities thereof. It should, of course, be ensured that no precipitation out of solution occurs under the recommended storage conditions and durations.

Due to the often very high concentration requirements for safety studies mentioned earlier, only a few compounds allow for the use of fully aqueous solution formulations across the entire dose range of a toxicology study. Other approaches to increase the solubility of a test compound and allow for the use of solution formulations include the addition of cosolvents, surfactants, or complexing agents (these approaches will be covered in a separate section in this chapter); however, they introduce the risk of vehicle effects, and care needs to be taken to stay within the concentration and dose volume limits for each excipient depending on the test species.

17.4.2 Suspension formulations

In cases where the solubility of the test compound is not sufficient to dissolve the dose into a solution formulation, a suspension formulation is typically developed. For BCS class I and III compounds, dosing of a suspension might still result in a dose-proportional increase in the exposure across most of the dose range, since the compound might have enough solubility/dissolution rate in the aqueous environment of the GI tract to fully dissolve and get absorbed, even if the solubility is not high enough to dissolve the entire dose in the dose volume administered. The advantage of a suspension formulation is that it typically uses a simple, aqueous vehicle, often a Methocel (hydroxypropyl methylcellulose) solution with a low level of surfactant to aid in the wetting of the solid test compound, so that the risk of vehicle effects is minimal. The suspension usually can be prepared by the stepwise addition of small portions of the vehicle to the entire amount of test compound in a mortar and thorough trituration with a pestle after each addition of vehicle to ensure the complete wetting of all solid particles and the formation of a fully homogeneous suspension. In cases where this trituration method is not feasible due to the large scale of the suspension preparation, the suspending vehicle can be added directly to the test compound and followed by more vigorous mixing to obtain a homogeneous suspension. However, the use of a suspension formulation also adds certain challenges and complexities to the formulation preparation, handling, and storage that do not exist for solution formulations. If during the toxicology study the dose volume is subdivided from a bulk suspension for each dose, the homogeneity of the suspension needs to be ensured during the entire use-time period. In addition, an appropriate procedure for resuspension after storage, such as sonication,

stirring, or shaking, needs to be identified to ensure that a homogeneous suspension can be obtained prior to every dose after the formulation was stored for a period of time. In cases where the test compound settles and forms a solid cake during storage that is difficult to break up and resuspend, continuous stirring during the entire use-time period might have to be employed. Changes in particle size of the suspended test compound during the use-time period also need to be evaluated. Particle size reduction caused by sonication or stirring during each resuspending cycle or especially during continuous stirring on storage has to be avoided. At the same time, particle size growth (Ostwald ripening) during repeated refrigeration/warming cycles over the course of the use-time period might become an issue. Lot-to-lot particle size and morphology variations of the test compound can potentially lead to batch-to-batch differences in homogeneity of the dosing suspension or its tendency to settle and cake during storage. Of course, particle size differences in a suspension can also lead to differences in the dissolution rate after administration and with variations in exposure, especially for BCS class II and IV compounds with low solubility. Lastly, the potential for crystalline form changes of the drug substance in a suspension needs to be investigated. For suspensions of amorphous drug substances or solid dispersions, there is a risk of drug substance crystallization from suspension over time.

17.5 FORMULATIONS FOR BCS CLASS II/IV COMPOUNDS USING ENABLING TECHNOLOGIES

Low aqueous solubility is the common attribute that classifies the BCS Class II/IV drug molecules. Compounds in these two classes have a tendency to exhibit low BA due to solubility-limited absorption or dissolution rate-limited absorption. Poor dissolution can be caused either by large particle size/small surface area or low solubility. To emphasize the importance of solubility, the dissolution/solubility-limited case is referred to as solubility-limited absorption, while the dissolution/particle size-limited case is called dissolution-limited absorption.[51] If the absorption barrier is dissolution/particle size-limited, then particle size reduction is a potential path forward. If the absorption is solubility-limited, improving both the solubility and the dissolution rate may be necessary.

For BCS II/IV compounds, extra formulation effort is often required in order to garner sufficient API exposures in vivo. This makes the formulation design more complicated compared to that for BCS Class I/III compounds. Theoretically speaking, if not limited by time and API quantity, an iterative trial-and-error type of comprehensive formulation screen may eventually result in a workable formulation. However, an aggressive timeline and the high cost of an API at the preclinical phase preclude a massive formulation screen. We advocate an efficient and rational formulation development process starting from the understanding of the root cause for the low solubility and dissolution rate.

The dissolution rate has been long established to be directly correlated to solubility, a thermodynamic property. The dissolution process can be viewed as a pseudochemical reaction, which is governed by thermodynamics. The Gibbs free energy, ΔG, is the balance between enthalpy, ΔH, and entropy, ΔS, at a given temperature T ($\Delta G = \Delta H - T\Delta S$). When a solute is introduced into a solvent system, dissolution will take place and continue until equilibrium is established or ΔG becomes zero, at which point the solubility of the solute is reached.[52] Since dissolution produces more randomness of molecules in the entire system (including solid and solution states), the ΔS is generally positive. Thus, dissolution is more driven by the net enthalpy of solution, ΔH_s.

The dissolving of a solid involves two consecutive processes: (1) the melting of the solute where the solute molecules break free from the solute-solute molecular interactions, described by the heat of fusion (ΔH_f^T) at the dissolution temperature, T, and (2) the mixing of the melt solute molecules with surrounding solvent molecules, namely, the establishment of solute-solvent intermolecular interactions, resulting in the change in heat of mixing (ΔH_{mix}). This two-step process can be expressed by the following equation[53]:

$$\Delta H_s = \Delta H_f^T + \Delta H_{mix} \tag{17.1}$$

From Eq. (17.1), it can be realized that in order to overcome the ΔH_s barrier during dissolution, we can either reduce the term of ΔH_f^T or facilitate the term of ΔH_{mix}. As such, Eq. (17.1) is the basis behind a wide array of solubilization and dissolution rate-enhancing techniques currently used by many researchers. For example, if the low aqueous solubility of a compound is due to its high hydrophobicity, indicated by a high log$P$ value, then we can change the solution environment to promote more solute-solvent intermolecular interactions. This can be done by adjusting the pH, adding a cosolvent, introducing lipids or surfactants, mixing with a complexing agent, or using combinations of these tactics. If the high heat of fusion (ie, high melting point) is the major barrier for achieving aqueous solubility, then techniques to break down the lattice energy are expected to be more effective. These techniques either introduce high-energy solid

forms such as amorphous or metastable solid forms or change the parent compound into a salt, if the compound is an acid or base strong enough to produce salts. The compounds with both high log$P$/log$D$ and a high melting point are apparently more difficult to formulate, and sometimes modification to the chemical structure is necessary to improve these properties.

It is also well-known that the dissolution rate correlates with the surface area of drug particles (refer to Nerst and Brunner equation in chapter: The Fundamentals of Diffusion and Dissolution of this book). Reducing the particle size, thus increasing the surface area, is probably one of the earliest approaches pharmaceutical scientists have used to increase BA.[54] The surface area term in the Nerst and Brunner equation is often viewed as the total surface of the exposed solid, such as the flat surfaces or the surface area of spherical particles. The initial dissolution is actually more driven by the corner and edge effect[55]; molecules at corners, edges, and defect spots are bound together with smaller interaction energies and are inclined to dissolve first. Particle size reduction technology creates proportionately more corners, edges, and other defects when the particles are turned into smaller ones, which help increase the dissolution rate.

When crystals are made into nanoparticles, the dissolution rate can increase significantly. The dissolution rate constant of a small particle, k_S, versus the dissolution rate constant of a large particle, k_L, can be calculated by Eq. (17.2)[55]:

$$\frac{k_S}{k_L} = \frac{C_{sol,S}}{C_{sol,L}} \times \left(\frac{r_L}{r_S}\right)^2 \tag{17.2}$$

where $C_{sol,S}$ is the solubility of the small particle, $C_{sol,L}$ is the solubility of the large particle, r_L is the radius of a large particle, and r_S is the radius of a small particle. The solubility ratio of a small particle over a large one can be calculated by the Freundlich equation (Eq. 17.3)[56]:

$$\ln\frac{C_{sol,S}}{C_{sol,L}} = \frac{2\gamma V_m}{rRT} \tag{17.3}$$

where γ is the interfacial tension, r is the radius of the small size particle, V_m is the molar volume of the particle (the molecular weight divided by the density), R is the ideal gas constant, and T is the absolute temperature. Let's use acetaminophen as an example to calculate the dissolution rate increase after particle size reduction to nanoparticles. Assuming the large particle has a particle size of 50 μm (a commonly used particle size for APIs), and the small particle is 800 nm, which is a good average size achievable by nanotechnology. With acetaminophen's molecular weight of 151.16 g/mol with a density of 1.2 g/cm^3 and the interfacial free energy of 30 mJ/m^2 at 310K (37°C),[56] the solubility of an $r = 800$ nm crystal is calculated to be 1.0037 times the solubility of a macrocrystal using Eq. (17.3). This shows that reducing the size to 800 nm only very slightly increases the solubility. The gain in the dissolution rate is mainly coming from the surface area increase: the dissolution rate ratio, k_s/k_L, of a $r = 800$ nm crystal versus a 50 μm crystal is $1.0037 \times (50{,}000\ \text{nm}/800\ \text{nm})^2 = 3920$, according to Eq. (17.2). Putting this value in practical perspective, if the large 50 μm particles take 60 min to dissolve, the 800 nm nanocrystals will take 0.9 s to dissolve, assuming discrete particles are dispersed and wetted in aqueous solutions. This represents potentially a tremendous dissolution rate advantage with the nanotechnology. It should be pointed out, however, no matter how fast the dissolution rate is, the highest concentration the nanocrystals can obtain is the crystalline solubility or slightly above the crystalline solubility. This contributes, at least partially, to the observation that nanotechnology is more suitable for BCS Class II compounds when the absorption is limited by the dissolution rate. With BCS Class IV (poorly permeable), formulations that generate supersaturation (higher concentration than equilibrium crystalline solubility) are expected to work much better. This aspect will be detailed in the context of developing in vitro dissolution method.

The aforementioned tactics that manipulate either the ΔH_f^T term or ΔH_{mix} term or particle size reduction have been collectively referred to as enabling technologies because they have enabled the delivery of challenging poorly water-soluble pharmaceutical compounds that otherwise would have been discarded during the drug discovery phase. All these approaches have been or can be used in the toxicology applications. Fig. 17.4 provides a systematic categorization of these formulations. Solution-based approaches are aimed at changing the solvent environment, while solid-based approaches reach the purpose by changing the solid properties of the API. Fig. 17.4 also emphasizes the importance of crystallization inhibition in every approach, since fast crystallization from supersaturated systems can reduce or even negate the effectiveness of these formulations. Nanosuspension has not become a mainstream formulation strategy for routine toxicological studies due to the challenges in scale-up. However, it is an excellent formulation for PK studies and early-stage toxicological studies. Readers are referred to published articles on this technology.[54,57,58] Complexing agents, such as cyclodextrins, can be effective solubilizers, but there are concerns of causing toxicological signals at large doses.[59–61] The other frequently used techniques are detailed in the following text with regard to mechanisms, application and methods of design, characterization, and preparation.

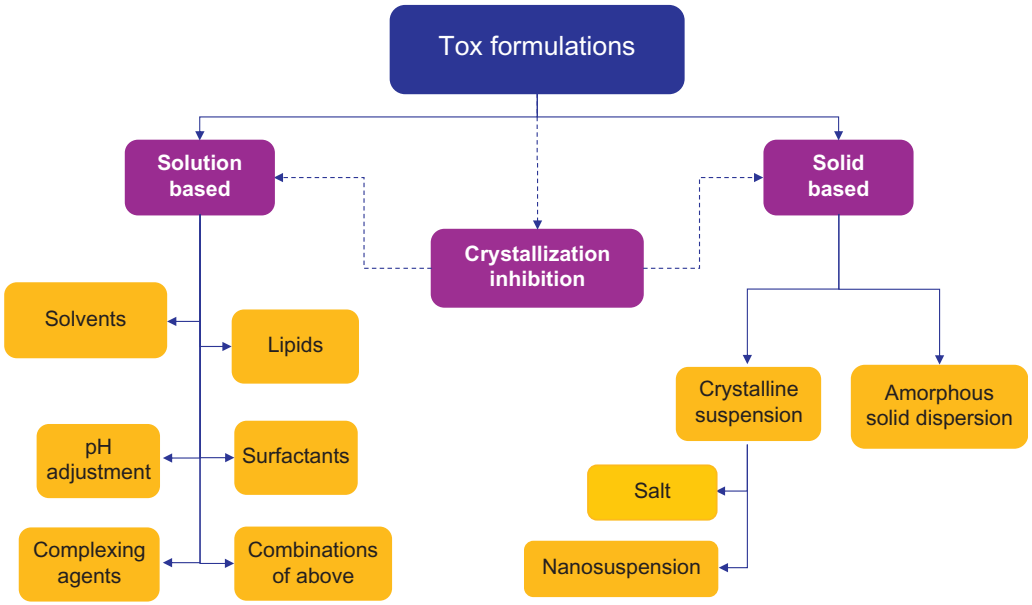


FIGURE 17.4 Categorization of enabling toxicology formulations for delivery of poorly water-soluble pharmaceutical compounds.

17.5.1 Solubilization by changing solution pH

As described in the previous section of factors that impact solubility, changing the pH is the most straightforward way to bring weakly basic or acidic pharmaceutical compounds into solution. The relationship between solubility and pH follows a set of defined models (equations) derived from Henderson–Hasselbalch ionization equations (reference to chapter: Solubility of Pharmaceutical Solids in this book). In other words, the solubility can be predicted if the intrinsic solubility and pK_a values of the compound are known.[62] Since pH is at log scale of [H^+], each pH unit change equals to a 10-fold [H^+] concentration change, and, therefore, one pH unit away from a pK_a will prompt a 10-fold solubility change (Illustrated in Fig. 17.2 and Eq. 17.4). For dibasic or diacidic compounds, the solubility will institute a 100-fold change with each pH unit change away from the second pK_a, and for tribasic or triacidic, a 1000-fold change. This makes pH adjustment a very powerful technique.

Using an acidic compound with a pK_a of 4.0 and an intrinsic solubility of 10 μg/mL as an example, based on one acidic pK_a—solubility model (Eq. 17.4), its total aqueous solubility (S_T) is 10 μg/mL at pH 2.0, 20 μg/mL at pH 4.0, 100 μg/mL at pH 5.0, 1000 μg/mL at pH 6.0, and 25,000 μg/mL at pH 7.4. If the goal is to make a 20 mg/mL solution, one can simply use a commonly available pH 7.4 phosphate buffer for this compound.

$$S_T = S_0\left(1 + 10^{\mathrm{pH}-\mathrm{p}K_a}\right) \tag{17.4}$$

In reality, oftentimes accurately measured intrinsic solubility and pK_a values are missing at the early stage. However, an estimation can be obtained from various sources, as described in previous section (Solubility Prediction and Screen). This information is used to assess the feasibility of the pH adjustment approach.

In practice, formulation starts with the targeted (or required) drug concentration, which is calculated based on the known targeted dose and the allowed dosing volume. For example, if the dose is 100 mg per kg of rat weight, then the target solution concentration is 10 mg/mL based on the commonly accepted dose volume of 10 mL/kg for an aqueous solution in a rat. To minimize API consumption in the prototype formulation screen, one could use the smallest possible volume, for example, by weighing about 1 mg and adding a final volume of 100 uL to make the final concentration of 10 mg/mL. To adjust the pH, the most commonly used base is NaOH and acid is HCl. They are added at a 1:1 molar ratio (API: pH modifier) or at an amount to reach a desired pH value. In the above case, if 10 mg/mL becomes a clear solution after pH adjustment, then the solubility target is successfully reached.

Although formulation preparation through pH adjustment is as simple as adding a counterion, many details still need to be taken into consideration and are listed in the following.

1. Chemical stability of the compound at extreme pH values. Chemical reactions such as hydrolysis are accelerated under extreme pH conditions.
   The degradant may exhibit increased solubility compared to the parent compound, which can be

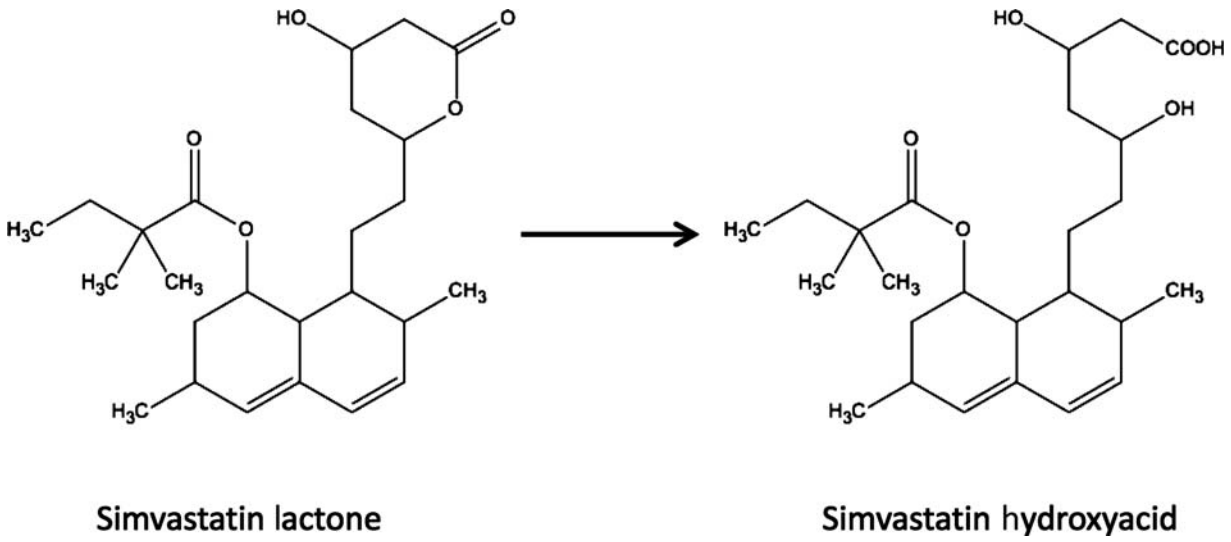


FIGURE 17.5 Under basic pH conditions, simvastatin lactone undergoes hydrolysis to simvastatin hydroxyacid, which becomes soluble at pH 13. *Source: Chemical Reaction adapted from Ungaro F, Giovino C, Catanzano O, Miro A, Mele A, Quaglia F, et al. Use of cyclodextrins as solubilizing agents for simvastatin: effect of hydroxypropyl-β-cyclodextrin on lactone/hydroxyacid aqueous equilibrium. Int J Pharm 2011;404:49–56.*

misleading if visual solubility is used to guide the direction of formulation development. One real life example is the solubility of simvastatin. The aqueous solubility of simvastatin was listed in the Merck Index[63] as 60 μg/mL at pH 1 and 70 mg/mL at pH 13, a greater than 1000-fold increase from pH 1 to pH 13. Simvastatin is a neutral compound, and its solubility at the pH 1−13 range should be a constant not a variant. It is impossible to observe the 1000-fold solubility difference across pH 1−13. What happened was that simvastatin is prone to hydrolysis at a basic pH, opening the ester bond in the six-membered lactone[64,65] to form hydroxyacid (Fig. 17.5). The degradant hydroxyacid is ionizable at a basic pH; thus it becomes much more soluble at pH 13, leading to an erroneous solubility assessment. Therefore, the chemical structure of the API should be carefully examined prior to pH adjustment, and stability should be particularly watched for when pH adjustment is to be attempted.

2. Solid form change during formulation preparation. At the early stage of discovery, compounds are purified by chromatographic separation and drying out the solvent. This often produces a metastable solid form (amorphous or higher-energy-state crystalline form), and conversion to a thermodynamically more stable form through solution-mediated transformation may occur at any time. This conversion process may be slow in neutral pH aqueous solutions due to low solubility. However, the pH adjustment creates a higher drug concentration that greatly increases the molecular collision rate to trigger nucleation and facilitate the solution-mediated phase transformation (SMPT) (refer to the section of transformation among solids in chapter: Crystalline and Amorphous Solids). Crystallization can occur through two ways: the transformation to a more stable crystalline form or to a salt with the added counterion. Either situation may lead to precipitation during the formulation preparation. If the precipitation occurs within several hours, one may still salvage the same formulation by preparing the formulation right before the dosing in a quick PK screen. For a long-term safety studies, reformulation will be necessary using a stable solid form.
3. Possible poor in vivo performance due to precipitation/crystallization as a result of pH gradient in GI tract. As described before, the solubility increases logarithmically with the pH decrease for basic compounds or with the pH increase for acidic compounds. Unfortunately, the solubility can also decrease with pH changes at the same dramatic level. This creates a large supersaturation to induce precipitation/crystallization. The solubility enhancement thus obtained may be lost in vivo due to pH gradient in the GI system. Using the same example acidic compound as illustrated before with a pK_a of 4.0 and an intrinsic solubility of 10 μg/mL, it can be formulated into a 20 mg/mL solution in a pH 7.4 buffer. When the solution formulation is dosed in rats (stomach pH of 4), up to 99.99% of the dose could potentially precipitate out in the stomach,

since the solubility at pH 4 is decreased to 20 μg/mL. Redissolving of the precipitates may not occur quickly enough to allow the dose to be absorbed, although the pH will gradually increase to 6.8 at which the solubility is higher. Similar phenomenon occurs to basic compounds, which dissolve better in the stomach, but then precipitate as the compound moves down the GI tract. It is important to know this precipitation tendency and profile before finalizing the formulation. An appropriate in vitro dual pH dilution test can be used for this purpose. Different additives such as crystallization inhibitors or solubilizers can be added to the formulation to maintain the supersaturation for a longer period of time or to reduce the severity of the precipitation. The effectiveness of the additives can be screened using the same in vitro dilution testing, which will be described in more details in the section Evaluating Formulation Performance by In Vitro Dissolution.

4. Safety concerns of extreme solution pH to animals. The safe pH range depends on the dosing volume, dosing duration (one dose versus repeated dose), and animal species. For example, a pH 2 solution is acceptable for one-time dosing but may cause irritation to the esophagus in long-term toxicological studies. The pH of Coca-Cola and Pepsi was reported as 2.5–2.6, which is probably at the lower end of a safe pH.
5. Limitation in dose loading. Every type of formulation approach has dose-loading limitations, although they are due to different reasons. The dose-loading limit by pH lies with the value of the intrinsic solubility and the acidity or basicity of the compound (pK_a). If the intrinsic solubility is in the ng/mL range, it is probably not possible to load the target high dose into a solution formulation, since changing the pH to 6 units away from the pK_a will only bring the solubility to the mg/mL range. On the plus side, dose loading through pH adjustment can be projected relatively easily, and the lengthy time to evaluate the feasibility can be avoided.

17.5.2 Formulation through suspension of salt

Salt formation is a well-established approach to overcome dissolution rate obstacles. In a solid dosage form for human use, salt is mixed with excipients, the blend of which is then tableted or encapsulated. This formulation approach is not usually adopted in preclinical studies. Solid dosage forms are not suitable for rodents, since they cannot swallow the big dose unit. For long-term studies in rodents, salt can be mixed into the food as a dosing route on the conditions that the taste is acceptable and adequate BA can be obtained. In reality, hardly any drug tastes good, especially when mixed into the food at high percentages. Gavage of freely flowing formulations (suspension) is usually the only dosing option for salt. Lager animals such as dogs can be trained to take pills; however, the tablets or capsules made for clinical studies are not able to deliver the large dose and/or lack the dose flexibility to meet the wide dose range that is often required in toxicological studies. These practical limitations determined why salt is usually dosed as suspensions in preclinical studies. This section, therefore, is focused on salt suspension formulations.

For salts of poorly water-soluble compounds, the major concern is the physical stability. Conversion back to the parent form can potentially occur when suspended into an aqueous medium. This is because the salt is a metastable solid form when the solution pH is away from the pH_{max}. An example of a pH-solubility profile of a weakly acidic drug, flurbiprofen,[66] is illustrated in Fig. 17.6, and used to explain the pH_{max} and its relevance to salt physical stability. Fig. 17.6 shows that when the solution pH is much less than the pK_a of flurbiprofen (ie, 2 pH units below the pK_a), the solubility of the drug is a constant, termed as intrinsic solubility (S_0). As the pH increases to $\geq pK_a$, the solubility continues to increase until the pH reaches the pH_{max} (at ~pH 7.3 for flurbiprofen) at which a salt is formed. At any pH above the pH_{max}, the solubility again becomes a constant, which equals to the solubility of the salt.

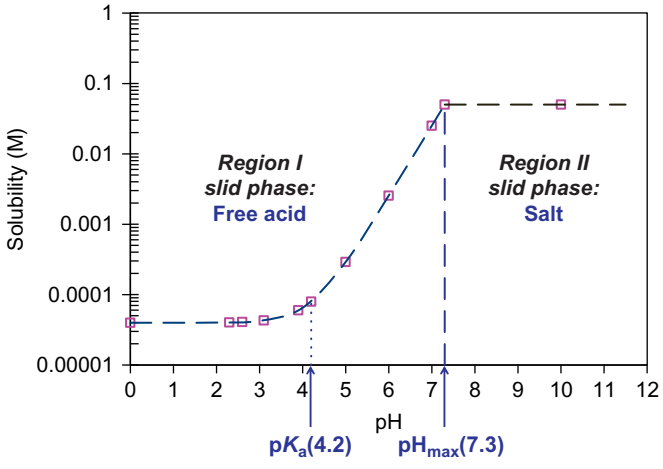


FIGURE 17.6 The pH-solubility profile of a weakly acidic drug, flurbiprofen, with a pK_a of 4.2. *Source: Adapted from Pudipeddi P, Serajuddin ATM, Grant DJW, Stahl PH. Solubility and dissolution of weak acids, bases, and salts. In: Stahl PH, Wermuth CG, editors. Handbook of pharmaceuitcal salts properties, selection, and use. New York, NY: Wiley-VCH; 2002. p. 19–41.*

Therefore, the pH-solubility profile is essentially a phase diagram. The pH_{max} divides the phase diagram in Fig. 17.6 into two regions: Region I is located below the pH_{max}, and Region II is located above the pH_{max}. In Region I, the free acid of flurbiprofen is the thermodynamically stable solid form, whereas in Region II, the salt of flurbiprofen is the thermodynamically stable form. At pH_{max}, the salt and the free form are at equilibrium with each other. In other words, the salt is a metastable form at any pH below its pH_{max}. If the salt particles of flurbiprofen are added to a solution with a $pH < 7.3$, the salt will eventually convert to the free acid of flurbiprofen. When equilibrium is reached, the solid recovered will be the free acid rather than the salt. This process is commonly known as disproportionation of a salt.[45,67] The same phenomenon exists for salts of basic compounds, where the pH-solubility relationship is a mirror image of that of an acid. The pH_{max} is located below the pK_a of the base, and the salt is metastable when placed in a solution with a pH greater than the pH_{max} and may convert back to the free base. The salt-to-parent conversion occurs through SMPT (refer to the section on transformation among solids in chapter: Crystalline and Amorphous Solids) where the salt dissolves first into the solution to reach a high enough supersaturation, which then triggers a nucleation of the crystal form of the free acid or free base.

To stabilize the salt in a suspension, we should consider both the thermodynamic and kinetic factors in the salt-to-parent form transformation process. Based on the pH_{max} theory illustrated in Fig. 17.6, if the suspension pH is below the pH_{max} of a salt of basic compounds, or above the pH_{max} of a salt of acidic compounds, the salt will be in a thermodynamically stable region and will never convert to less soluble parent form crystals. In reality, this is less likely to be achieved, as shown by Eq. (17.5) and Eq. (17.6). Eq. (17.5) is the Henderson–Hasselbalch equation of a monobasic drug. $[B]$ is the concentration of the unionized base, and $[B^+]$ is the concentration of the ionized base. Eq. (17.6) is written as the Henderson–Hasselbalch equation at the salt solubility equilibrium point,[68] where $[B]_s$ is the solubility of the unionized base, K_{sp} is the solubility product of the salt, and $\sqrt{K_{sp}}$ is the salt solubility, which can also be written as $[B^+]_s$. In a salt suspension, the maximum concentration of the salt is the salt solubility. Practically, before the salt solubility can be reached, the free form may have already crystallized if the nucleation of the free form is facile. For the sake of this discussion, let's assume the salt solubility can be reached quickly during the suspension preparation step. At the solubility of the salt, the solution pH will be at the pH_{max} of the salt, or the maximum pH of the salt suspension will be equal to pH_{max}. The pH is not able to go below the pH_{max} of the basic drug. Similarly, the pH of a salt suspension of an acidic drug will not reach above the pH_{max} of the salt.

$$pH = pK_a + \log\frac{[B]}{[B^+]} \tag{17.5}$$

$$pH_{max} = pK_a + \log\frac{[B]_s}{\sqrt{K_{sp}}} = pK_a + \log\frac{[B]_s}{[B^+]_s} \tag{17.6}$$

The discussion indicates that the salt suspension will not be thermodynamically stable, although being close to the pH_{max} does reduce the driving force for nucleation of the parent form crystals. In order to be within the thermodynamically stable region, an extra amount of counterion needs to be present, for example, adding NaOH to a sodium salt suspension until the pH reaches above pH_{max}. This is a viable approach, since a small pH increase above pH_{max} for basic compound salts or a small pH decrease to below pH_{max} for acidic compound salts will suffice to make the salts thermodynamically stable, as long as the final pH does not extend to extreme acidity or basicity to render safety concerns.

If the thermodynamic stability pH region cannot be reached due to safety concerns, the kinetics of the salt-to-parent form conversion in suspensions can be intervened using formulation and dosing techniques. Different methods can be tried to slow down the conversion process. The suspending agents, usually polymers used at 0.2–2%, result in a higher viscosity, which may slow down the conversion rate. Therefore, a prudent approach is to select a polymer or a combination of polymers that can act as both a good suspending agent and a strong crystallization inhibitor capable of preventing or significantly delaying the solid form transformation. In recent years, understanding of the mechanisms of crystallization inhibition by polymers has significantly improved from the studies of amorphous solid dispersion formulations.[69–76] The methodology in selecting the polymers will be discussed in more detail in the amorphous solid dispersion section. An alternate way was reported to dose salt particles suspended in an organic solvent.[77] It was thought that the lower solubility of the salt in the solvent would delay the conversion kinetics. The solvent-based dosing method introduces an unnecessary potential safety risk to animals and removes the benefit of the salt suspension as a benign aqueous formulation. Stabilizing the suspension through a polymer-based methodology is a more reasonable approach.

The length of physical stability required for the salt suspension depends on the nature of the study and the different dosing routines established in the different study units. It is desirable and sometimes necessary to obtain a suspension formulation with 1 week or

longer physical stability, since a weekly formulation preparation schedule is preferred for resource savings. In the worst-case scenario when week-long physical stability cannot be obtained, there is an option to prepare the salt suspension on daily basis or right before dosing.

Prior to making salt suspensions, the suspension density (weight percentage) is calculated based on the dose and the dosing volume. For example, if the dose is 300 mg/kg of rat at a dosing volume of 10 mL/kg, the minimal suspension density is 30 mg/mL (or 3%). The procedure for preparing a salt suspension is the same as that described previously (section Suspension Formulations) for BCS Class I/III compounds. It should be mindful, however, to use the initial higher solid-to-liquid ratio (higher surface area), which helps drive the salt to quickly reach salt solubility to avoid or minimize the salt-to-parent form conversion. The acceptable threshold of the suspension viscosity can be tested by pulling the suspensions into a syringe with the dosing gavage tube attached and expelling the suspension out of the syringe. If the loading and unloading of the suspension from the syringe is smooth and does not require excessive efforts, feasibility of dosing can be assumed. Essentially, the highest possible dose is limited by the suspension viscosity.

After the suspension is administered in vivo, the absorption is dependent on the dissolution rate from the salt particles. API particle size and distribution should be controlled for reproducible performance. The particle size control should consider not only the particle size of the dry API but also the actual API particle size in the suspension. An aggregation or glomeration of particles in a suspension due to wetting problems reduces the effective surface area. A wetting agent (eg, 0.1% Tween 80) is typically applied to overcome wetting issues; however, in this case, caution should be exercised for the potential of the wetting agents (ie, surfactants) to promote crystallization of the parent form.[78] Particle size distribution control also should consider minimizing the Ostwald ripening effect where particles grow at the expense of smaller and more soluble particles. Polymers find themselves yet another application here to inhibit or slow down the crystal growth.[79,80] Appropriate and smaller particle size also helps stabilize the suspension without fast sedimentation and pass the suspension through the narrow diameter of the gavage tube used in dosing small animals (rat or mouse).

It should be noted that salt selection is a prerequisite for a successful salt suspension formulation to use in preclinical studies. Salt selection is often a part of the solid form selection process, which is complex and involves multiple functional areas from discovery, process chemistry, and formulation development. Salt selection should consider the dissolution rate, stability, manufacturability, and formulatability of the salt. Physical stability is one of the important criteria. We often have a tendency to identify a salt with a higher solubility and dissolution rate; however, a salt with a higher solubility produces a pH_{max} more far away from the neutral biological pH, which leads to a higher risk of salt-to-free form conversion (physical instability) when the salt is placed in a neutral pH aqueous medium.[68] Selecting a physically more stable salt will greatly reduce the challenges in formulating and dosing the salt suspension for preclinical studies.

17.5.3 Solubilization through cosolvents

Cosolvent-based formulations are widely used due to relatively fast formulation development and the generally straightforward preparation for daily use. The useful solvents are those semipolar such as PEG 300, PEG 400, propylene glycol (PG), ethanol, DMA, and so forth. These solvents possess the ability to form hydrogen bonding with water molecules and induce a certain degree of polarity in nonpolar drug molecules, thus bringing about miscibility between nonpolar drug molecules and water.[81] For this reason, semipolar solvents are quite versatile. They can be added to aqueous solutions to produce a simple aqueous/solvent blend or an aqueous/multiple solvent blend, or they can be added to a lipid/surfactant to formulate more complex lipid-based formulations (to be separately described). Some of them can be used as a neat solvent under certain restrictions.

The solubility behavior of drug molecules in solvent-aqueous mixtures is complex. Fig. 17.7 schematically describes two types of solubility profiles. In solvent system A, the solubility of a drug molecule increases initially with an increasing percentage of a good solubilization solvent but plateaus at certain solvent percentage, after which the solubility goes downhill with further increase of the good solvent.[82] In solvent system B, however, the solubility keeps increasing with the addition of more good solvent, and the maximum solubility resides at the 100% good solvent composition.

The solubility increase with a good solvent follows a semilogarithmic relationship established by Yalkowsky et al. (refer to Eq. 1.58). When more than one solvent is used in the mixture, a similar relationship applies,[22] as described in Eq. (17.7). In Eq. (17.7), $[S_{tot}]$ represents the overall drug solubility, which is the sum of the intrinsic solubility, $[S_0]$ and the solubility created by each solvent. The magnitude of the solubility gain by a solvent is the product of the volume fraction of the solvent, f_i and the solubility power of the solvent, σ_i, which is a unique property of the solvent. Eq. (17.7)

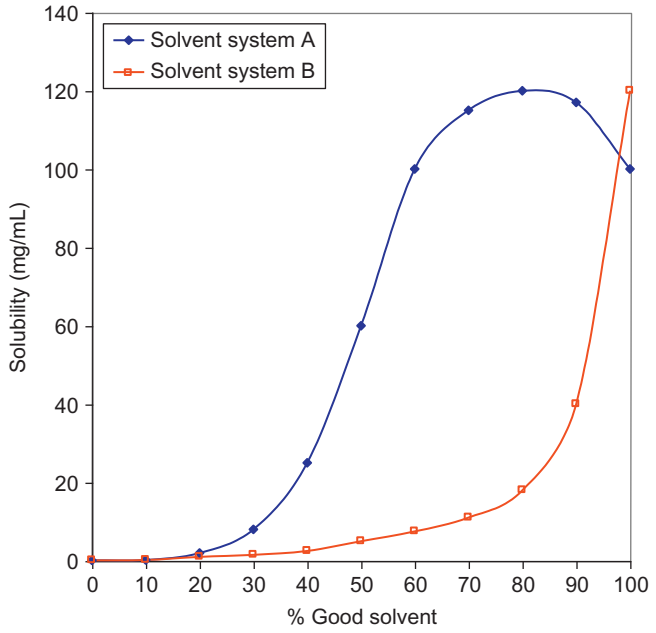


FIGURE 17.7 Schematic illustration of solubility profiles in mixtures of solvent and aqueous solutions.

provides a theoretical basis that multiple solvents can be used together, since the solubilization power is additive.

$$\log[S_{tot}] = \log[S_0] + \sum_{i=1}^{n} \sigma_i \cdot f_i \qquad (17.7)$$

The formulation using cosolvents starts with a solubility screen in various cosolvents. The goal is to identify the best solvent or solvent mixtures for further formulating into a blend with water or aqueous buffer. A solubility screen can also be done in predetermined blends of solvents and buffers that have been working for a certain group of compounds. For example, if a group of compounds historically dissolve well in PEG 400 and have a basic pK_a of ≥ 6, a formulation blend with PEG 400 and a slightly acidic buffer (ie, pH 4 buffer) serves a good starting point.

Cosolvents often provide an excellent solubility increase over the aqueous solubility. Unfortunately, cosolvents can be toxic, especially when used neat or at a high percentage (refer to Table 17.3). Another issue is that, if mixed with aqueous buffer in the final formulation, the solubility power decreases at a semilog scale, which threatens to achieve target drug loading. This dilemma can be mitigated to some degree using several modifications, for example, using a mixture of multiple solvents. With an appropriate blending ratio of various solvents, it is possible to increase the overall solvent content to maximize the solubility while maintaining or reducing the vehicle toxicity. This is based on the assumption that the toxicity enacted by each solvent is additive, which may not be true in every case. Additional consultation with toxicologists in the preclinical safety area is encouraged. Another approach is to add pH adjustment simultaneously. Li and Zhao[22] determined that the pH/cosolvent combination works synergistically to significantly increase the solubility. As previously described, the same caution on chemical stability and side effect at an extreme pH must be exerted. It should also be noted that pK_a values change in the presence of a solvent; a base will become a weaker base, and an acid will become a weaker acid. Therefore, the pH of the cosolvent solution may change to a more extreme pH when the same amount of counterion as used in the pure aqueous formulations is added. A final pH measurement is recommended before dosing in the animals. An alternative way is to mix the solvent with a buffer at a safe pH. To further increase the solubility, the aqueous phase can contain other solubilizers such as surfactants or cyclodextrins. Finally, all of the methods can be combined if necessary. For example, for an acidic compound (ie, pK_a of ~ 4) with a good solubility in PEG 400, ethanol, and a chemical structural element to potentially going into the cyclodextrin cavity, a mixture of 40% of PEG 400, 10% ethanol, and 50% of 30% beta-cyclodextrin in a pH 7.4 buffer may provide a desirable outcome. However, some solubilization mechanisms may not compatible with each. In such cases, it is wise not to combine these approaches.

Special attention should be paid when the formulation contains a high-solvent content. On one aspect, diluting a high-solvent-content formulation by aqueous medium tends to induce fast precipitation based on the steep portion of the solubility-phase diagram (Fig. 17.7). Dilution is an inevitable event when dosing in vivo. This risk should be evaluated by an appropriate in vitro dilution test, which should be an integral part of the formulation development. The rational design of an in vitro dilution test will be covered in the section of Evaluating Formulation Performance by In Vitro Dissolution. Another aspect comes from the safety perspective. The dosing volume of a high-solvent-content formulation may have to be reduced for safety reasons. The allowable dosing volume is often different among different preclinical species (see Table 17.2). Therefore, the safe dosing volume should be determined before rushing into the formulation effort in order to avoid the last minute surprise.

17.5.4 Lipid and surfactant-based formulations

Lipid and surfactant-based drug delivery systems are among the most versatile and effective formulation

approaches to enhance the BA of poorly water-soluble compounds. The number of commercial formulations, however, is limited mainly due to practical limitations in drug loading (thus a high pill burden), the concerns on API and excipients chemical stability, the requirement of compatibility with capsule shells, and a more complex manufacturing process. Most of these disadvantages will vanish in the preclinical setting where dosing volume can be large, much more flexible, long-term chemical stability is not required, and the capsule is not a component of the formulation. Therefore, this category of formulations is well suited for use in preclinical studies. However, lipid and surfactant-based formulations are often viewed as more mysterious. One reason is probably due to the difficulties in predicting the solubility in lipid excipients as opposed to other systems such as pH adjustment, cosolvent addition, or complexation. Another reason is because the performance of the formulations is less predictable by conventional in vitro dissolution or dilution testing. Adding more to the mystery is the large variety of lipid excipients; the numerous commercial names of lipids and surfactants alone may overwhelm a novice formulator. Therefore, the successful development of lipid and surfactant-based formulations will require the learning of rational selection of the excipients and the use of appropriate in vitro dissolution methods to determine the potential in vivo performance.

17.5.4.1 Commercially available excipients

The first and foremost step in learning lipid and surfactant-based formulation is to gain knowledge in the various excipients. The ability to formulate well is heavily dependent on the understanding of excipient properties and miscibility among the excipients. The list of commercial available excipients is shown in Table 17.5. This table separates the excipients into three categories: water-soluble, water-insoluble, and surfactants. Such a division is helpful for users to conveniently locate the needed excipients based on the class of lipid formulations to formulate. Phospholipids listed under the water-soluble group actually have limited solubility in water and can be moved to the water-insoluble category. The water-insoluble group is the most varied and diverse group, encompassing natural oils and fatty acids; semisynthetic monoglycerides, diglycerides, and triglycerides; semisynthetic PEG derivatives of glycerides and fatty acids; polyglyceryl fatty acid esters; cholesterol; and phospholipids. Each of these subcategories further branches out through varying fatty acid chain lengths, degrees of esterification (mono-, di-, or tri-), and compositions. Some of them are unique blends named as particular trade names such as Gelucire marketed by Gattefosse. To make it more complex, different vendors assign different trade names to their products; this gives rise to a similar blend being called different names. Fortunately, comprehensive summaries of the water-insoluble group have been written in a book chapter by L. Gibson.[83] Readers are encouraged to read the reference to gain more comprehensive knowledge.

17.5.4.2 Selection of excipients

Before selecting the excipients, it is beneficial to understand the structural constituents of lipid-based formulations. Pouton and Porter[85] divided lipid-based drug delivery systems into four types:

- Type I: Oils (triglycerides, diglycerides, and monoglycerides) only.
- Type II: Oils and less water-soluble surfactants.
- Type III: Oils, water-soluble surfactants, and cosolvents.
- Type IV: Water-soluble surfactants and cosolvents (no oils).

Compounds from modern drug discovery mostly do not solubilize well in pure oils, especially not in long-chain glycerides. Achieving a high drug load in Type I and in some instances in Type II is relatively difficult. On the other hand, these compounds tend to solubilize better in polar oils, phospholipids, surfactants, and cosolvents that are part of Type III/IV formulations. In addition, Type III and IV are SEDDS or self-microemulsifying drug delivery systems. They are easily dispersible in aqueous medium, leading to high supersaturation level, which is the major reason for the significant BA enhancement. These properties propelled the much more frequent use of Type III and Type IV than Type I and II in preclinical applications.

When the lipid-based formulation is diluted into an aqueous GI milieu, the drug molecules are protected from precipitation by remaining as a solution in lipid droplets, surfactant micelles, and/or other different lipid/surfactant aggregates. This indicates that the solubility in the excipients is an important factor not only for achieving high drug loading in the formulation but also for enhanced in vivo performance. Therefore, the logical selection of excipients starts with a solubility screen in the excipients. In general, the excipients that provide the highest solubility are the ones that should be considered. Water-miscible solvents such as PEG 400 and ethanol provide high drug loading owing to their unparalleled solubilization power and their ability to reduce the viscosity of the formulations for easier handling; therefore, they are indispensable in formulating Type III and IV formulations. As described before, pure solvent-based formulations often result in massive precipitation when diluted into aqueous

TABLE 17.5 List of Commercially Available Excipients

Water-Soluble	Water-Insoluble	Surfactants
Dimethylacetamide (DMA)	Beeswax	Polyoxyl 35 castor oil (Cremophor EL)
Dimethyl sulfoxide (DMSO)	Oleic acid	Polyoxyl 40 hydrogenated castor oil (Cremophor RH 40)
Ethanol	Soy fatty acids	Polyoxyl 60 hydrogenated castor oil (Cremophor RH 60)
Glycerin	D-alpha-tocopherol (vitamin E)	Polysorbate 20 (Tween 20)
N-methyl-2-pyrrolidone (NMP)	Corn oil monoglycerides, diglycerides, triglycerides	Polysorbate 80 (Tween 80)
PEG 300	Medium chain (C8/C10) mono- and diglycerides	D-alpha-tocopheryl polyethylene glycol 1000 succinate (TPGS)
PEG 400	Long-chain triglycerides	Solutol HS-15
Poloxamer 407	Castor oil	Sorbitan monooleate (Span 20)
Propylene glycol	Corn oil	PEG 300 caprylic/capric glycerides (Softigen 767)
Hydroxypropyl-β-cyclodextrin	Cottonseed oil	PEG 400 caprylic/capric glycerides (Labrasol)
Sulfobutylether-β-cyclodextrin (Captisol)	Olive oil	PEG 300 oleic glycerides (Labrafil M-1944CS)
Phospholipids	Peanut oil	PEG 300 linoleic glycerides (Labrafil M-2125CS)
Hydrogenated soy phosphatidylcholine (HSPC)	Peppermint oil	Polyoxyethylene 8 stearate (PEG 400 monostearate)
Distearoyl phosphatidylglycerol (DSPG)	Safflower oil	D-alpha-tocopheryl polyethylene glycol 1000 succinate (TPGS)
L-a-dimyristoylphosphatidyl choline (DMPC)	Sesame oil	Polyoxyl 40 stearate (PEG 1750 monostearate)
L-a-dimyristoylphosphatidyl glycerol (DMPG)	Soybean oil	Peppermint oil
	Hydrogenated soybean oil	
	Hydrogenated vegetable oils	
	Medium-chain triglycerides	
	Caprylic/capric triglycerides derived from coconut oil or palm see oil	

Adapted from literature Strickley RG. Solubilizing excipients in oral and injectable formulations. Pharm Res 2004;21:201–30.[84]

solutions. However, when combined with lipids and/or surfactants, the precipitation will be buffered by the lipids and surfactants. Ionic surfactants such as sodium dodecyl sulfate have limited solubility in lipids; therefore, nonionic surfactants, especially those with high hydrophile-lipophile balance values such as Cremophor EL, Cremophor RH40, Tween 20 and vitamin E TPGS, have been shown to be compatible with lipids. Surfactants play interesting roles in Type III and IV formulations. When in the formulation blend, surfactants sometimes are antisolvent, since they may not solubilize the API as effectively as lipids and cosolvents. When the formulation is diluted into an aqueous media, surfactants stimulate the formation of emulsion droplets and also form micelles to solubilize the API. Because of the dual roles, the selection of surfactants should consider the drug solubility in neat surfactants as well as in surfactant micelles in the presence of an aqueous medium. The former is related to drug loading in the formulation, and the latter is related to formulation performance in vivo.

As shown in Table 17.5, there exist large numbers of excipients, and screening all of them for solubility is apparently not feasible with limited API quantity and time. Several ways can be considered to narrow the screen funnel. One of them is to take into account the physicochemical properties of the API. For example, moderately lipophilic compounds (log*P* in the 2–4 range) are usually not very soluble in long-chain triglycerides. Polar compounds are more compatible with polar lipids such as phospholipids and oleic acid. Another is based on the fact that lipid excipients that do not rely on lipolysis as a prerequisite for drug release are preferred for rapid drug absorption.[86] This will eliminate the long-chain glycerides and put emphasis on the medium-chain monoglycerides, fatty acids, and monoesters of fatty acids. Last but not least is to examine the properties of

the lipids. Behind the façade of commercial names, the vast number of lipids is actually classified into smaller sets of substances as shown in Table 17.5; for example free fatty acids; short-chain monoglycerides, diglycerides, and triglycerides; medium-chain monoglycerides, diglycerides, and triglycerides; propylene glycol esters; and phospholipids. Choosing one lipid from each class will form a representative lipid panel. A visual solubility screen in this panel of lipids is conducted to generate a solubility profile based on which more focused screen can then be pursued in a smaller subset of lipids. Similar approaches can be applied to the screening of the surfactants and solvents. A solubility screen funnel provides an efficient way to quickly identify useful excipients.

17.5.4.3 Formulation development

After the promising excipients are identified from the solubility screen, which may include lipids, solvents, and nonionic surfactants, the next step is to formulate vehicles by combining these excipients at appropriate proportions.

The first question in every new formulator's mind is: What is the appropriate proportion to mix these excipients identified through the initial solubility screen? In other words, how to determine the percentage of a surfactant and/or a cosolvent to add to the oil (in Type III), or how much cosolvent should be added to a surfactant (in Type IV)? The answer to this complex question is surprisingly simple in concept. It depends on the miscibility (forming a one-phase or isotropic system) among those excipients. The determination of the miscibility is a straightforward experiment simply by mixing the excipients at different ratios and then recording whether or not a single-phase solution is obtained. However, it is tedious and time consuming. The amount of surfactant required can be determined by adding incrementally increased amounts of surfactant to each lipid, evaluating the stability of the emulsions formed upon dilution, and then selecting the best emulsified vehicle. This is also a tedious process. Since the API after being dissolved into the vehicle may change the phase diagram as well as the emulsification outcome, determining precisely the blank vehicle's miscibility and its ability to emulsify seems not a wise use of time. In a fast-paced preclinical setting where the formulation is needed right away, these tedious and time-draining experiments often could not be afforded. A useful shortcut is to roughly determine the miscibility by starting from a center percentage point, for example, at 50% lipid with 30–40% solvent and 10–20% surfactant. If the first blend is not quite successful, then a couple of adjustments at 10% incremental changes can be made while keeping the surfactant at the highest possible level to ensure the best self-emulsifying ability. This can quickly close into the miscible region and identify the initial prototype formulations for first PK study. Further optimization in the composition of the prototype formulations can be continued after receiving promising PK results. Clearly, prior knowledge in the miscibility of the selected excipients, either from publication or personal accounts, provides valuable information to pinpoint the miscible blends. The accumulation of the knowledge comes with diligence and attention to detail. Every time a clear vehicle is observed, record the composition, and save the information. These miscible blends can be further investigated with interest to identify potential trends within the excipients with similar physicochemical properties. This can quickly expand the miscibility database that has been gradually established with each formulation endeavor. In turn, the invaluable experiences will expedite the next formulation design and development.

The next step following the successful development of the prototype excipient blends is to measure the final solubility of the API in each of the blends, usually by using a visual solubility method. Although the solubility is often additive in nature from the individual excipient, the solubility needs to be confirmed in the final blend. Interaction among the excipients themselves is not uncommon, which may influence the API solubility in the vehicle blend. Due to the viscosity of the vehicle, the visual solubility is routinely measured under moderate heat (ie, 40°C). Hence, it should be cautious to make sure the solubility can be maintained when the formulation is cooled to room temperature in order to determine the drug loading more accurately.

The solubility in the selected blends and the allowed dosing volume are the first criterion to triage in different prototype formulations. Those with high enough solubility to deliver the target dose are preferred. As described in the solvent-based systems, safety is related to the dosing volume. Similar to the solvent selection discussed previously, the allowable dosing volume of lipids, surfactants, and solvents is often different among the different preclinical species. Reliable excipient safety data, usually available from manufacturers' GLP toxicological or in-house studies, should be consulted prior to making a decision. An adjustment in the formulation composition may be needed in order to meet the safety requirements.

17.5.4.4 Formulation characterization and selection

Before advancing the selected prototype formulations to animal PK and BA studies, in vitro characterization of the formulation should be carried out. First, the physical stability should be monitored. The high solution

concentration provides a high-driving force for polymorphic transformation. A more stable polymorph of the API crystal or a crystalline adduct with any of the excipients may materialize over time, either of which leads to the precipitation of the otherwise clear solution. Once the polymorphic transformation has taken place, the drug loading will have to be redetermined. Lipid excipient adduct is often more detrimental than API polymorphic transformation because the lipid as a dominating component in the formulation exists at high thermodynamic activity, greatly favoring the adduct formation and decreasing the solubility. These adduct-formed formulations are often discarded. In addition, the chemical stability of an API can be a challenge for formulation development. The stability required usually coincides with the formulation preparation intervals on the dosing site. If the formulation is to be prepared weekly, then week-long stability is required. If heat is used to solubilize the API, then its effect on the stability should be included in the stability study. The last and the most critical characterization of the lipid/surfactant-based formulations is the dissolution/dilution testing. When diluted into an aqueous medium, the formulation should quickly disperse and form fine droplets and should not precipitate massively as crystalline particles. Because of the high importance of the in vitro dissolution, a special section is dedicated to how to develop a suitable in vitro dissolution method for enabling formulations. Details can be learned following the section of Evaluating Formulation Performance by In Vitro Dissolution.

17.5.5 Amorphous solid dispersions

Amorphous solid dispersion (ASD) renders the API into amorphous material, which is dispersed into a matrix consisting of either 100% polymer (single or multiple polymers) or polymer mixed with other excipients such as surfactants and lipids.

ASD formulation comprises of two unique properties that have collectively contributed to the significantly increased BA for poorly water-soluble drugs. The first property comes from the amorphous API. As described before, by disrupting the crystal lattice of the crystalline drug substance, the resultant higher-energy amorphous form provides a higher apparent solubility having a potential for higher dissolution rate. Thermodynamically, the higher dissolution rate is given; however, kinetically, the potential of higher dissolution rate oftentimes cannot be fully realized within the GI residence time if the amorphous drug is not properly formulated. Dosing a neat amorphous solid, which is analogous to a 100% drug loading ASD, may give a mediocre performance even when there is no crystallization. Thus, the second ASD property attributed to the success is the formulation itself, which is equally if not more important than the first property. Appropriately selected excipients in ASD formulations can drive the amorphous API to quickly dissolve and reach amorphous solubility in a solution, along the way creating nanosized amorphous particles. Ideally, these excipients then maintain this amorphous solubility by preventing or minimizing crystallization within the time scale relevant to in vivo absorption. ASD formulations have been utilized to deliver many life-saving drugs commercially.[87,88] Numerous papers and recently two books have been directed to amorphous pharmaceuticals and ASDs.[89–92] The readers are referred to these references. The focus here is on preclinical application of ASDs.

In preclinical studies, ASDs are found to be more readily applicable than in commercial human use. One of the greatest challenges in human formulation is the attainment of high drug loading for fewer or smaller pills. When increasing the drug loading, the API dissolution rate from the ASDs is inevitably decreased in addition to a potentially worsened physical stability. In preclinical ASD formulations, it is much more flexible to incorporate a wider variety and a higher percentage of hydrophilic excipients (ie, large amount of surfactants) to improve the dissolution rate at high drug loads. These excipients used at high percentage in preclinical applications are often not allowed in human formulations either for safety concerns or for processing reasons during commercial dosage form manufacturing. In addition, extra excipients such as binders or lubricants necessary for tableting or encapsulation can be eliminated, which further contributes to higher drug loading.

Another major challenge for commercial ASD formulation is the maintenance of physical stability during the product shelf life. This again is less demanding for preclinical applications due to the shorter duration of these studies, and in cases when the physical and chemical stability is not adequate, storage of ASD under refrigerated conditions during the relatively short period can be done to circumvent the shortcoming. Further, ASDs are more acceptable than lipid/surfactant/solvent-based formulations in that they are less toxic due to the nature of essentially being an aqueous suspension. Therefore, ASDs have become increasingly popular in preclinical studies.

The formulation development process for preclinical ASDs follows a similar path to that for human ASD formulations. The difference is that some steps in the process can be bypassed to accommodate the tighter timeline imposed at the preclinical stage, which will be discussed.

The formulation development starts with the evaluation of the physicochemical properties of the API. It has been believed that the amorphous solubility advantage is the prerequisite to justify for the amorphous solid dispersion approach. The amorphous solubility advantage can be either calculated or measured. The calculation involves the approximation of the free energy difference (ΔG) between the amorphous and the crystalline forms using the Hoffman equation[93]:

$$\frac{\sigma_a}{\sigma_c} = e^{\frac{\Delta G}{RT}} = e^{\frac{\left(\frac{\Delta H_f \cdot \Delta T}{T_m}\right)\left(\frac{T}{T_m}\right)}{RT}} \tag{17.8}$$

where $\frac{\sigma_a}{\sigma_c}$ is the ratio of amorphous-to-crystalline solubility, ΔH_f is the molar heat of fusion of the API crystal, R is the ideal gas constant, T is the temperature at which the solubility is concerned, T_m is the melting point, and ΔT is the temperature difference between the melting point and the solubility temperature. To improve the accuracy of the estimation, the solubility ratio should be further corrected, as described by Bogner and coworkers,[94] for the extent of ionization of the amorphous form versus the crystalline form at the pH of interest and for the amount of water the amorphous form absorbs, which reduces the thermodynamic activity of the amorphous solid. This correction requires the pK_a value of the API and the moisture sorption profile of the amorphous form.

The solubility difference can also be directly obtained through the solubility measurement if both crystalline and amorphous forms are available and the amorphous form is physically stable in aqueous media during solubility measurement. Either the calculation or the experimental measurement requires a reasonably pure and crystalline API, which is often lacking at the start of preclinical studies in addition to the time that entails.

In reality, accurate information on the amorphous advantage is nice to have but not essential when making a decision to proceed with an ASD. The reasons are: (1) in most cases, the amorphous form is more soluble than the crystalline counterpart; (2) as described, another attribute to determine the success of an ASD lies in the formulation. The hydrophilic matrix has been shown to enable liquid–liquid phase separation (LLPS) or nanoparticle formation during dissolution.[71,95–98] The extent of the dissolution rate increase by LLPS particles may well exceed the dissolution rate gained from the amorphous solubility advantage. The notion that only drugs with a high melting point can benefit from ASD formulation is not universally true; and (3) at the preclinical formulation stage, an ASD is not a first-line approach. It is pursued when other simpler approaches (such as a simple solvent formulation or pH adjustment) have failed. ASDs are often evaluated in parallel with lipid-based formulations. In other words, an ASD is competing with the lipid formulation approach, not with a crystalline form formulation. Therefore, the time and effort used to obtain an accurate amorphous advantage (ratio of amorphous solubility over crystalline solubility) can be saved for the moment until approaching the clinical formulation development stage.

A much more useful property to determine at this stage is the crystallization tendency of the API, which sheds light on the formulation strategy and design. Crystallization tendency is holistically a function of configurational entropy and molecular mobility.[99,100] Configurational entropy is intrinsic to the molecular structure. Simpler chemical structure and smaller molecular weight are attributed to lower configurational entropy. Lower configurational entropy means a higher probability to assume the proper orientation and conformation for crystal nucleation. Higher molecular mobility is related to smaller molecular size and volume and also is influenced by environmental conditions (viscosity and temperature). Qualitatively, crystallization tendency can be ranked by the timing of crystallization when a supersatured solution is created, when heating/cooling the melt in a differential scanning calorimetry (DSC),[101] or by stress testing of an amorphous solid made from solvent evaporation.[102] The DSC method[101] classifies drug molecules into three categories: Class I is a faster crystallizer (crystallization occurs when the melt is cooled to an ambient temperature), Class II is a medium-fast crystallizer (crystallization occurs when the cooled melt is reheated), and Class III is a slow crystallizer (no crystallization occurs when the cooled melt is reheated). Different formulation strategies should be adopted for compounds with different crystallization tendencies.

17.5.5.1 *Formulation for fast crystallizers*

For fast crystallizers, stabilizing the amorphous form is evidently the major challenge. The key to success will be to select the right polymer to effectively inhibit crystallization. Polymers play an important role in preventing the amorphous material from recrystallization via crystallization inhibition mechanisms.[71,72,74,75,103] The requirement for the polymer is twofold. First, the identified polymer should be capable of inhibiting crystallization in a solid state for achieving an adequate shelf life. Second, the polymer should be able to inhibit crystallization during the dissolution (in solution state) to maintain supersaturation for a superior in vivo performance. In preclinical studies, the latter is shown to be decisively more important, since ASDs are dosed as aqueous suspensions. Commonly used polymers include hypromellose acetate succinate (HPMCAS), hydroxypropyl methylcellulose (HPMC), copovidone,

Soluplus, polyvinylpyrrolidone (PVP), hydroxypropyl cellulose (HPC), and Eudragit.

The selection of a polymer can be done through a screen process. The crystallization inhibition in a solution should be the first to be addressed, as it is the most challenging to success. Based on a recent molecular dynamics simulation study, hydrophobic and van der Waals interactions are the dominating forces in establishing the drug-polymer interaction in aqueous solutions.[104] This indicates that polymers capable of forming strong hydrophobic interactions are more effective in inhibiting the crystallization of poorly water-soluble and hydrophobic drugs in an aqueous media (such as GI fluid). Polymers with both hydrophobic and hydrophilic moieties in the structure, such as HMPC or HPMCAS, are found as better choices.[105]

In solution-state polymer screen experiments, different polymers are predissolved into aqueous solutions (pH 1 and pH 6.8 buffers or SSIFs) at concentrations achievable in vivo. Supersaturation in these aqueous solutions is created by adding a high drug concentration stock solution made in an organic solvent such as methanol or by adding a high drug concentration solution made at extreme pH values. The extent and duration of supersaturation as well as the rate of crystallization from the solution are then monitored by a suitable analytical technique or multiple techniques depending on needs. For example, the drug concentration can be assayed by UV/Vis or HPLC. The crystallization can be detected by PLM, PXRD, or Raman spectroscopy. The polymer that maintains the highest supersaturation without crystallization is preferred. This screen scheme can also be used to investigate the combinations of polymers that may provide stronger effect than a singular polymer. For example, a polymeric blend of copovidone and HPMC was found to have a superadditive effect to prolong the supersaturation.[106] Some researchers have applied a more sophisticated HTS to execute the polymer screen in solutions.[107] This allows for more quickly mapping out the useful formulation space by evaluating more polymers and more supersaturation levels in a shorter time.

The second step in the polymer selection is to screen the polymers that can maintain amorphous stability in a solid state. Different from the solution state where crystallization inhibition is mainly mediated through hydrophobic and van der Waals interactions, solid-state-specific interactions such as H-bonding[108] or ionic interaction[109] can play a significant role. Polymers with a high H-bond donor or acceptor propensity are the candidates for consideration if the API structure has the matching H-bond acceptors or donors. However, it should be noted that water is ubiquitous. The hydrophilic polymers used in ASDs will retain or pick up water during formulation preparation and storage. For example, griseofulvin-PVP K30 solid dispersion crystalized upon exposure to moisture during storage due to PVP losing its inhibition capacity.[110] Whenever there exist water and hydrophobic molecules, there will be hydrophobic interactions taking place. Therefore, the ability of the polymer to engage in hydrophobic interaction with the drug is still important in solid-state crystallization inhibition. The set of polymers selected based on crystallization inhibition in an aqueous solution is presumably also useful in preventing crystallization in the solid state when exposed to moisture.

The screen for polymers to work in the solid state is done differently than in the solution state. Small-scale ASDs with different polymers each at different API loadings are prepared, such as by casting films,[111,112] and subjected to accelerated thermal and/or moisture stress conditions with miscibility or lack of crystallization being monitored periodically. Miscibility usually refers to a kinetically stable state where the API exhibits the apparent solubility in the polymer[113] without phase separation. Drug-polymer miscibility can be measured experimentally using various techniques such as Raman spectroscopy,[114] solid-state nuclear magnetic resonance (NMR)[115], or a combination of atomic force microscopy-based infrared, thermal, and mechanical analysis.[116] If phase separation occurs, the drug-rich phase becomes supersaturated, which may lead to crystallization. The detection of crystallization can be made using PLM or PXRD, which in general is easier to do than determining miscibility. Through this process, polymers that achieve miscibility or lack of crystallization at the highest drug loading can be determined and rank-ordered. As stated previously, toxicological studies are usually shorter in duration. If the formulation does not crystallize under ambient temperature and humidity for several months, the physical stability is deemed acceptable. The ideal situation is to identify a polymer or a combination of polymers that can stabilize amorphous material in both solution and solid state. In some challenging cases, it is sometimes useful to use the right polymer to tailor an appropriate drug-release rate at which the degree of supersaturation during dissolution is controlled to minimize the driving force for nucleation.[117] Since some polymers perform better in crystallization inhibition and others do better in enhancing dissolution rate, sometimes a combination of more than one polymer can be used to optimize the formulation.[118–120]

Increasing drug loading for the fast-crystallization class is especially challenging. Achieving a higher drug loading above 30–50% is often questionable.

The use of a surfactant for improving the dissolution rate should proceed with caution, since surfactants may increase nucleation and crystal growth rate.[78] For extremely fast crystallizers, selecting a more soluble polymer is perhaps the only effective means to increase drug loading. This is because if the API portion is solubilized in the polymer, the ASD system is thermodynamically stable and will not crystallize. Although the Flory–Huggins interaction parameter-based solubility prediction method is widely reported,[121,122] experimentally measured solubility values are more accurate and reliable. Determining the thermodynamic solubility in polymers using melting point depression-based methods have been reported.[123–126] API and polymer mixtures with an increasing percentage of API were cryomilled, annealed at different temperatures, and analyzed by differential scanning calorimetry (DSC) until undissolved crystals remain, and thus the upper and lower bounds of the equilibrium solubility temperature were determined.[124] The procedure required many DSC runs, which may take days to complete. This method was shortened significantly by taking an empirical approach utilizing a combination of differential scanning calorimetry measurements and a reliable mathematical algorithm to construct the solubility curve of a drug in a polymer.[126] A polymer screen through this solubility measurement is lengthy and costly, but it may identify a niche to formulate the extremely fast crystallizer. Finally, lower the drug loading may be still possible to formulate extremely fast crystallizers in ASDs.

17.5.5.2 Formulation for slow crystallizers

Slow crystallizers are ideal compounds for ASD formulations. Since crystallization is a small risk, it opens up the possibility to formulate higher-drug-loading ASDs. Different from faster crystallizers, the major challenge is switched from how to stabilize the amorphous form to how to enhance the drug release. Therefore, a different formulation strategy is considered.

Slow crystallizers that are poorly water-soluble are often hydrophobic. In spite of incorporating hydrophilic polymers, at some threshold of increasing drug loading depending on the hydrophobicity of the API, the solid dispersion will become too hydrophobic to dissolve into the aqueous phase at a desirable rate. The improvement in the dissolution rate will rely on selecting the appropriate polymers and secondary excipients.

The formulation development also starts with the polymer screen, albeit with a different focus. The miscibility or solubility in the polymer is still important, since any phase-separated hydrophobic drug region will hinder the dissolution, while the molecularly dispersed hydrophobic drug molecules may dissolve together with the fast-dissolving carrier molecules. Like in the case of fast crystallizers, miscibility is depending on polymer types. For example, it was shown that the miscibility of telaprevir is increased from 10% in HPMC or copovidone to 30% in HPMCAS.[116] Thus, the polymer screen is based on the dissolution profiles of different polymers incorporated into different drug loading ASDs made at small scale. Polymers that meet the criteria of rendering a fast dissolution rate and stimulating nanosized particle formation should be ranked as top selections.

Problematic dissolution rates from high-drug-loading ASDs can also be rescued by adding extra highly water-soluble excipients. These secondary excipients consist of mainly surfactants and in some cases lipids. The surfactants listed in Table 17.5 can all be potentially utilized in preclinical ASDs. As mentioned previously, there is a freedom of intercalating the high percentage of these secondary excipients (ie, 10–30%) for preclinical amorphous solid dispersion formulations as long as the T_g is above the ambient temperature, and the excipient level is tolerable by the animals at the intended dose. Physically, the ASDs containing a large amount of the secondary excipients may exist in multiple phases rather than in a single phase. The coexisting of several phases seems not having any negative impact to either the physical stability or the reproducibility of in vivo performance, especially for the slow crystallizers. Rather it brings the benefit of achieving higher drug loading while attaining a fast drug dissolution rate. Increasing the drug loading greatly reduces the dosing volume and subsequently the animal dosing time every day. Therefore, effort should be made to formulate the drug loading as high as possible. In vitro dissolution is the appropriate tool to gauge whether or not the dissolution rate is fast enough, which will be discussed in the section of Evaluating Formulation Performance by In Vitro Dissolution.

Evidently, formulation strategy for medium-fast crystallizers will be between the fast and slow crystallizers.

17.5.5.3 Preparation of prototype amorphous solid dispersion formulations

After the polymers are identified from the screen experiments, the next step is to prepare the different prototype ASD formulations, which are then characterized to arrive at the final lead formulation. To recapitulate, for the fast crystallizers, these selected polymers should be promising to inhibit crystallization in the solid state as well as in the solution state to achieve the desirable drug-polymer miscibility or solubility, while for the slow crystallizers, the best polymers are those that can promote high drug loading and at the

same time obtain an acceptable drug dissolution rate. Practically, what determines the formulation composition of various prototype ASDs is the drug loading. The drug-loading evaluation has been built into part of the polymer and the secondary excipients screen as delineated.

There are different methods available to prepare ASD formulations.[127] In order to make small-scale prototype formulations, the most convenient method is solvent evaporation using a rotary evaporator. This technique is commonly used in chemical synthesis, and a rotary evaporator is readily available commercially. In cases where formulations consist of a high percentage of cellulose-type polymers (ie, HPMCAS or HPMC without surfactants), rotary evaporation is difficult to operate due to film formation, which causes difficulties to recover the solid. The hardship in the recovery of the solid in general is tolerable because of the small scale.

A key to successful solvent evaporation is to locate a suitable solvent that meets the following criteria: (1) sufficient solubility for the API. The higher the solubility, the less solvent to consume; (2) sufficient solubility for the polymer and other excipients in the formulation; (3) low boiling point so that it can be easily removed; and (4) meeting the residual solvent safety requirement per ICH guidelines.[128,129] Solvents are divided into three classes per International Council for Harmonisation (ICH) guidelines: Class 1 solvents contain known human carcinogens and should be avoided in all; Class 2 solvents can be used but should be limited to permitted daily dose (PDE) due to nongenotoxic animal carcinogens; Class 3 solvents are preferred since they are relatively safe with low toxic potential, and the PDE is 50 mg or more per day. The list of Class 3 and Class 2 solvents and their PDE values can be found from the ICH guidelines. The most preferable solvents include Class 3 solvents, such as acetone, ethanol, and ethyl acetate, and Class 2 solvents, such as methanol. A combination of solvents may present an advantage to increase the solubility for API and excipients and lower the boiling point (after forming azeotropes).

After the solvent is found, the API and the polymer at the ratio of designed drug loading are dissolved to form a clear solution, sometimes with the help of moderate heating. The clear solution is then dried under vacuum in a rotary evaporator, usually heated slightly below the boiling point of the solvent. A second drying is carried out in a vacuum oven to further reduce the solvent to the acceptable level. The dried mass is recovered and ground to fine particles. These amorphous solid dispersion powders should be stored away from light and moisture prior to characterization and use.

17.5.5.4 *Characterization of prototype amorphous solid dispersion formulations*

The studies designed to rank the initial prototype amorphous solid dispersions are summarized into four categories (Fig. 17.8). The first three are universal for characterizing amorphous formulation systems, while the last is unique to animal studies.

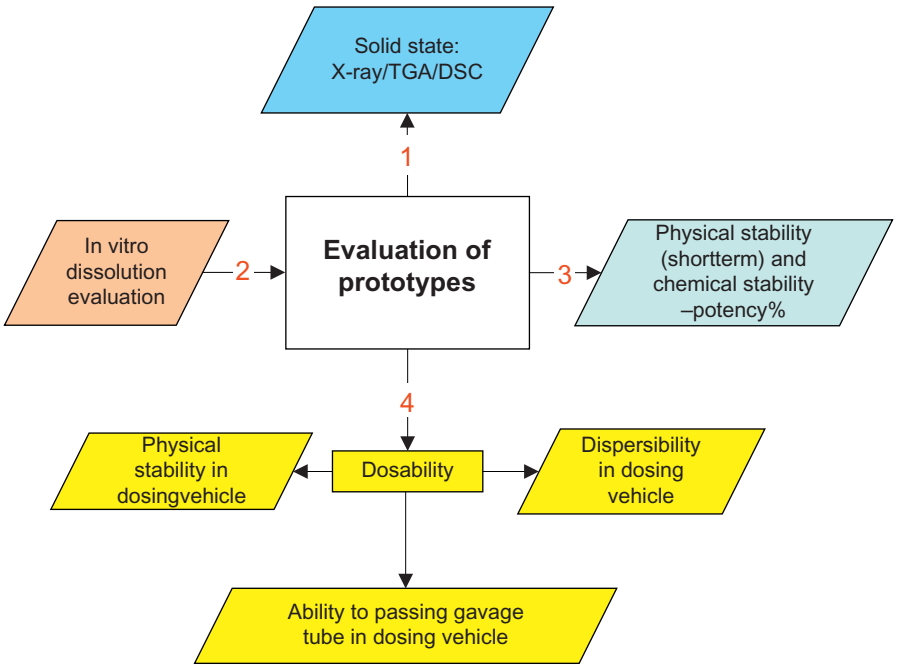


FIGURE 17.8 Characterization of initial prototype amorphous solid dispersion formulations.

The number one priority is to make sure the API is indeed turned into an amorphous state (tested by PXRD or PLM), the organic solvent level is meeting ICH criteria (determined by TGA or gas chromatography), and the glass transition temperature (tested by DSC) is preferably significantly higher than the ambient temperature. The second is to confirm that the formulation is doing what it is supposed to do: increasing the dissolution rate and maintaining supersaturation. It is very important to select the right dissolution method to assess this potential. This is so critical that a special section is dedicated to the in vitro dissolution method (see the section Evaluating Formulation Performance by In Vitro Dissolution). The third area of testing is to ensure the formulation is physically stable within a reasonable time frame. Since long-term physical stability is not required, stress testing at 40°C/75% RH or 65°C/25% RH is an overkill. Several months of stability in a closed container at ambient conditions is adequate. More extensive evaluation may be necessary for faster crystallizers. Thermal chemical stability against rotary evaporation temperature and secondary drying temperature, on the other hand, will be important in order to make it at a larger scale for long-term toxicological studies.

The fourth evaluation is specific to dosing in animals. The word dosability is created to describe the ability of any ASD formulation to be dosed smoothly. Similar to dosing the salt, gavage of freely flowing formulations (suspension) is usually the only option for administering an ASD to rodents and sometimes even for dogs when the dose cannot be carried within the limited numbers of capsules. The dosability as a suspension can be jeopardized in several situations: (1) if the amorphous form is quickly crystallized during the suspension preparation step, which basically obliterates the advantage of the amorphous solid dispersion; (2) if the powders do not disperse well into a smooth and uniform suspension or produce a suspension too viscous to load into syringes; (3) if foaming is substantial, which is a common problem when surfactants are part of the formulation; and (4) if the particles of the suspension are too large to pass through the narrow diameter gavage tube. Therefore, each of these four risks is to be rigorously evaluated prior to finalizing the formulation. Some of them are less challenging to resolve. An antifoam agent can be added to quench the foam. The particle size can be reduced by more aggressive milling. The growing of particle size in a suspension can be controlled using polymers. The viscosity can be reduced by selecting less swelling polymers.

However, the most intractable problem is the crystallization. In the presence of an abundance of water, the molecular mobility is as high as it can be, which potentially induces fast crystallization. The polymers contained in the formulation may lose the battle when against the strong headwind. Slowing down the drug release and/or finding a better suspending solution are possible ways to alter the kinetics of the crystallization. For example, when using an enteric polymer, such as HPMCAS or HPMCP as the polymeric carrier in ASD formulations, the amount of API (lopinavir) leaching out was shown to be below the amorphous solubility when the ASD is suspended in a lower pH buffer (pH 3 or 5).[130] This may help prolong the stability in some cases. A suspending solution containing different and more effective crystallization inhibitors (polymer or combination of polymer and surfactant) may also improve stability. A time window of several hours is needed from suspension preparation to dosing. If this dosing window cannot be met, an ASD is not a viable formulation approach.

Based on the results from the four areas of characterizations, the prototype formulations are compared, triaged, and narrowed down to the final or the final a few. The most promising formulations will be advanced to in vivo PK studies, often done in a dose-escalating fashion in comparison with other formulations such as lipid-based formulations. If the ASD formulation demonstrates advantages over the others, it will be declared as the final formulation for future studies. At this time, how to scale-up the ASD preparation will be the next discussion topic.

17.5.5.5 Scale-up the amorphous solid dispersion formulations

Larger scale ASD preparations can be done mainly through three technologies: rotary evaporation, spray drying,[131] and hot-melt extrusion.[132] Each technique has its unique pros and cons. The selection of the method is dictated by the polymers selected in the formulation, the availability of the technology, the know-how in-house, the thermal stability of the API, and the nature of the preclinical study. For example, if the formulation uses dominantly HPMC or HPMCAS-type polymers, spray drying is necessary. However, the fine and fluffy particles from spay drying create more work to prepare the suspensions. Hot-melt extrusion is solvent-free and produces high-density particles for easier suspension preparation, but it generally takes a longer time and more API to develop in addition to the need of specific equipment. The rotary evaporation is a more suitable technique for making an ASD supply for the preclinical studies for noncellulose polymers.

The principle for large-scale rotary evaporation is no different from a small scale. The API and formulation excipients are all dissolved into the solvent of

choice to form a solution, which is then heated under vacuum to remove the solvent. However, the same procedure is much more difficult to execute considering the scale. Assuming the solubility of the API is 50 mg/mL in the solvent, the calculated volume of the solvent required to turn 1 kg of API into ASD will be 20 L. It will take several batches to deliver the 1 kg API even using a large 10 L flask. Applying sufficient vacuum during the evaporation of 10–20 L of solvent will require a very powerful vacuum pump. The rate of solvent removal is a parameter to control in cases of faster crystallizers; a slow evaporation rate may cause crystallization. After finishing the evaporation, the massive amount of ASD cake is scraped out from the flask and milled into fine powders, which are then subjected to a second drying. The major challenge faced during the second drying process is the removal of the residual organic solvent. Much higher drying temperatures and longer drying times than the lab scale are often required to dry several kilograms of ASD powders to below the residual solvent specifications. The extensive heating may cause chemical degradation and crystallization. The drying temperature, if above the Tg of the formulation, may transform the ASD into a rubbery state and fuse the particles together. These risk factors should be evaluated in a small-scale experiment prior to the scale-up. For example, small ASD samples are heated in the anticipated drying temperature range over hours and days and periodically tested for chemical stability and particle morphology. The maximum drying temperature and time are then defined as control parameters during the scale-up. The final product is subjected to an analytical analysis in potency, crystallinity, and stability at GLP regulations if intended for GLP toxicology studies. The material is also tested for dissolution and dosability to ensure the larger scale product does not change these properties as compared with the small-scale prototype that was initially dosed in animals. These analyses define the formulation storage condition and shelf life for the intended toxicology studies.

17.6 EVALUATING FORMULATION PERFORMANCE BY IN VITRO DISSOLUTION

The theory of drug dissolution has been covered extensively in this book (refer to dissolution chapters in this book). This section will focus on evaluating the enabling formulations that create high supersaturation of the API in aqueous media, including pH adjusted, cosolvent, lipid and surfactant-based formulations (particularly Type III and IV), salts, and amorphous solid dispersions.

Since the mechanism of the enabling formulations is via creating supersaturation in the GI milieu, it is palpable that the performance of these formulations will count on the maintenance of the supersaturation. It quickly became clear that the standard dissolution test conducted under sink conditions is not able to evaluate these formulations because supersaturation is depleted under the sink conditions. Efforts have been, therefore, directed to imitate the nonsink conditions in vivo with the objective to predict in vivo performance. For example, the physiological conditions that can influence the supersaturation have been built into the design of an in vitro simple dual-pH dilution model for rats.[133] As illustrated in Fig. 17.9, the rat model mimicked how a formulation sample experiences the pH/medium change from a pH 4 buffer in stomach to a pH 6.5 FaSSIF in small intestine, the

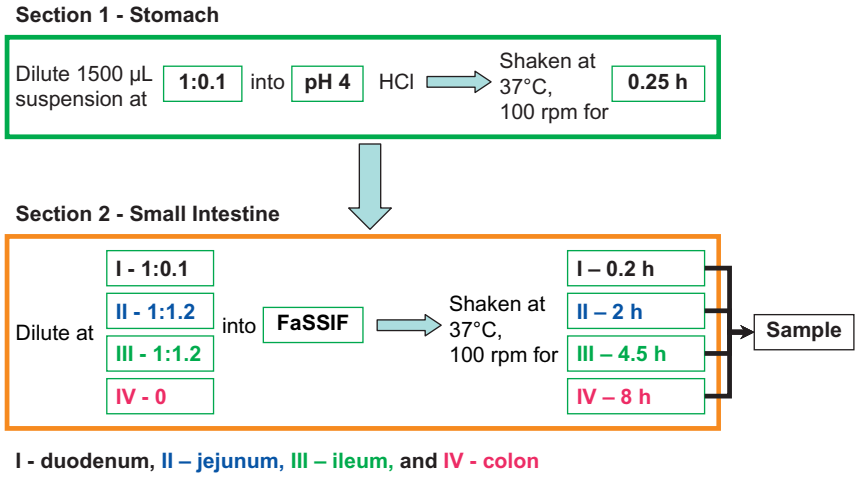


FIGURE 17.9 Schematic of a simple dual pH-dilution method to mimic dilution/dissolution in rats for evaluating enabling formulations.[133]

concentration change due to gradual dilutions along the GI tract, and different residence times at each section of the GI tract. Samples were taken from 0.2 to 8 h after the sample reached the intestine compartments and centrifuged to separate the supernatant from the solid phase. The solution phase was analyzed by HPLC assay to obtain the supersatuation concentrations. The formulation that resulted in the highest supersaturation was determined as the best formulation. This method was found to rank a series of formulations into a correct order in line with in vivo data for the model compound investigated.[133] This dissolution test can be conducted in a small beaker or vial in lieu of the standard US Pharmacopeial (USP) apparatus with ≥ 500 mL of dissolution medium. The simple dual-pH dilution is a straightforward experiment but requires hands-on manipulations to change the dissolution medium, add diluent, and take and analyze samples. Using the concept of dual pH and dilution was also reported in a microdissolution pH-shift test,[134] where the drug was introduced into a pH 2 solution for 20 min, followed by dilution into FaSSIF for another 180 min. The sample analysis was made automatic using commercially available μDISS Profiler (Pion Inc., Billerica, Massachusetts) fitted with UV-fiber optic probes for monitoring the drug concentration in situ. The UV dip probe offers an advantage of collecting the rich information of the drug concentration over a long period of time in a hands-free fashion, yet the interference from the background solution has to be removed or minimized before reliable data can be generated.

In spite of the effort to make the in vitro dissolution test more biorelevant, reasonable prediction of the formulation performance in vivo has been mostly elusive. Increasing the drug concentration by adding solubilizing agents such as cyclodextrin, surfactants, or solvents appeared to confer reduced permeability, as shown in rat jejunal permeability of several poorly water-soluble model compounds.[135–137] Some studies showed that monitoring the drug concentration in a water-immiscible 1-octanol layer added to the top of the aqueous dissolution medium (biphasic dissolution) may mimic the in vivo sink conditions and provide a better understanding of the supersaturation maintenance and absorption process.[138–140] The biphasic method started to link supersaturation with absorption but with limitations. The method simplified the permeation process for lacking a physical membrane barrier between the aqueous and 1-octanol layers and for treating the solubility in 1-octanol as the dominating factor for permeability.

To design a meaningful in vitro dissolution, it is essential to understand how the supersaturation is generated by the enabling formulations and what exactly the supersaturating system is undertaking to drive drug absorption. The former question has been made clear for a while: creating supersaturation by enabling formulation is analogues to launching a parachute into the sky.[141] However, the latter question was not clearly addressed. A further question was: When the parachute is launched toward the sky, how high can it go? These important unknown questions have been tackled in recent years by a series of publications.

In 2011, Alonzo et al.[93] reported that an felodipine amorphous solid dispersion at a low drug loading of 10% in a 90% HPMC polymer rapidly dissolved into a pH 6.8 buffer, forming a semitransparent and colloid suspension. The tiny droplets in the buffer were determined to be nanosized particles by dynamic light scattering, and the felodipine concentration in the colloid suspension was found to equate with its amorphous solubility. The colloidal species could also be generated by solvent switching and pH shifting. The transformation from an initially clear solution to a colloid phase was found to occur after a highly supersaturated solution was generated and can be attributed to a physical phenomenon called LLPS.[96] In 2013, Raina et al.[142] further confirmed that the LLPS threshold occurs at the amorphous solubility by studying several poorly water-soluble compounds. The data implied that the highest supersaturation any formulation can reach is the amorphous solubility above which the system separates into a two-phase system, wherein one phase is an initially nanodimensioned and drug-rich phase, and the other is a drug-lean continuous aqueous phase.

More interestingly, Raina et al. in the same paper[142] investigated the impact of LLPS formation on passive membrane permeability using a side-by-side diffusion cell. The flux, measured as a drug diffusional rate from the donor chamber to the receiver chamber in the diffusion cell, is theoretically correlated to the solute thermodynamic activity or free drug concentration. As shown in Fig. 17.10, when the drug supersaturation level was increased in the donor side, the flux of drug diffusion into the donor side increased linearly but abruptly stopped right at the LLPS threshold, above which the flux remained as constant. This indicates that the highest driving force for membrane permeation is when the solute activity is at amorphous solubility or the LLPS point. A higher concentration than LLPS is short lived, and the system quickly equilibrates to the LLPS level.

In a closely followed paper, Raina et al.[143] further demonstrated that the commonly used solubilizing agents, when added at excess to above the solubility of the solute, decreased the solute activity and consequently decreased the flux for membrane permeation. However, once more drug is added to reach

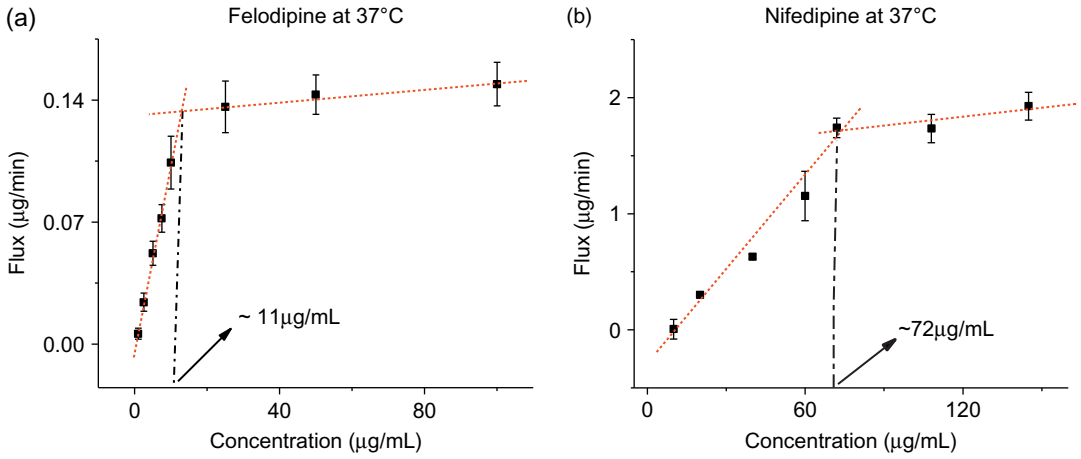


FIGURE 17.10 Diffusive flux versus concentration profiles for (a) felodipine and (b) nifedipine at 37°C. The flux of felodipine and nifedipine increases linearly with concentration up to the amorphous solubility (LLPS onset concentration). Further increases in concentration resulted in a plateau. *Source: Adapted from Raina SA, Zhang GGZ, Alonzo DE, Wu J, Zhu D, Catron ND, et al. Enhancements and limits in drug membrane transport using supersaturated solutions of poorly water-soluble drugs. J Pharm Sci 2014;103:2736–48.*

solubility, the flux is recovered to a normal level. This paper cleared the confusion about the role of solubilizing excipients and provided guidance on how to appropriately utilizing these excipients in enabling formulations.

These findings have had profound implications to formulation design as well as to in vitro dissolution method revamp. LLPS occurs at phase separation between the drug-rich phase and drug-lean phase (aqueous phase saturated with the drug). The numerous nanosized drug droplets are in equilibrium with the aqueous phase. Once the free drug molecules initially dissolved in the aqueous phase are absorbed, the nanosized drug-rich phase rapidly releases more drug to replenish the lost drug molecules in solution, thus maintaining the highest solute thermodynamic activity for absorption.[144] As long as LLPS can be maintained within the GI absorption window, the formulation is or is close to optimal. This new understanding in the role of supersaturation and LLPS is the foundation for designing an LLPS-based in vitro dissolution method that will effectively monitor the generation and maintenance of the highest solute (drug) thermodynamic activity.

Based on the core value of LLPS, a simple in vitro dissolution method can be designed to fit the purpose. Fig. 17.11 schematically describes the proposed LLPS-based in vitro dissolution method for evaluating the enabling formulation using a rat model in comparison with a human model. The scheme can be easily adapted to different animal models by replacing the in vivo parameters correspondingly.

As shown in Fig. 17.11, the LLPS concept-derived new dissolution method still takes into account the biorelevant conditions: (1) applying dual pH if the compound or excipients have pK_a value(s) within the physiological pH range of 1–7.4, (2) mimicking stomach and small intestine residence time and potential dilution along the GI tract, (3) using simulated small intestine fluid (FaSSIF or FeSSIF) as a dissolution medium, and (4) testing at in vivo temperature (37°C). However, the sample analysis is redesigned with the focus on determining whether or not the LLPS is generated and maintained. Nanosized droplets are manifested as a semitransparent or colloidal suspension, which can be monitored visually. A microscope (PLM) and/or nano particle size analyzer can be used to monitor the phase behavior and LLPS particles; the precipitates should be amorphous in nature and preferably in nano dimensions. The smaller the precipitated particles are, the better the formulation will be. Minimally, there should not be massive precipitation to crystalline form within several hours.

The free drug concentration determination in the sample analysis scheme presented in Fig. 17.11 requires interpretation. The drug concentration level measured to the solution phase of the samples should not be taken at face value, since part of the solution concentration can be due to the solubilizing agent, which is not free drug. An accurate measurement of free drug concentration can be accomplished using membrane diffusion-type experiments (ie, side-by-side diffusion cell[142]). It may not be feasible in the preclinical stage to engage sufficient time and resources in designing and conducting an appropriate membrane diffusion-based test. If the diffusion test is not affordable, the free drug concentration can be estimated by comparing the apparent concentration obtained from the dissolution samples with the measured amorphous solubility obtained under corresponding placebo conditions, as

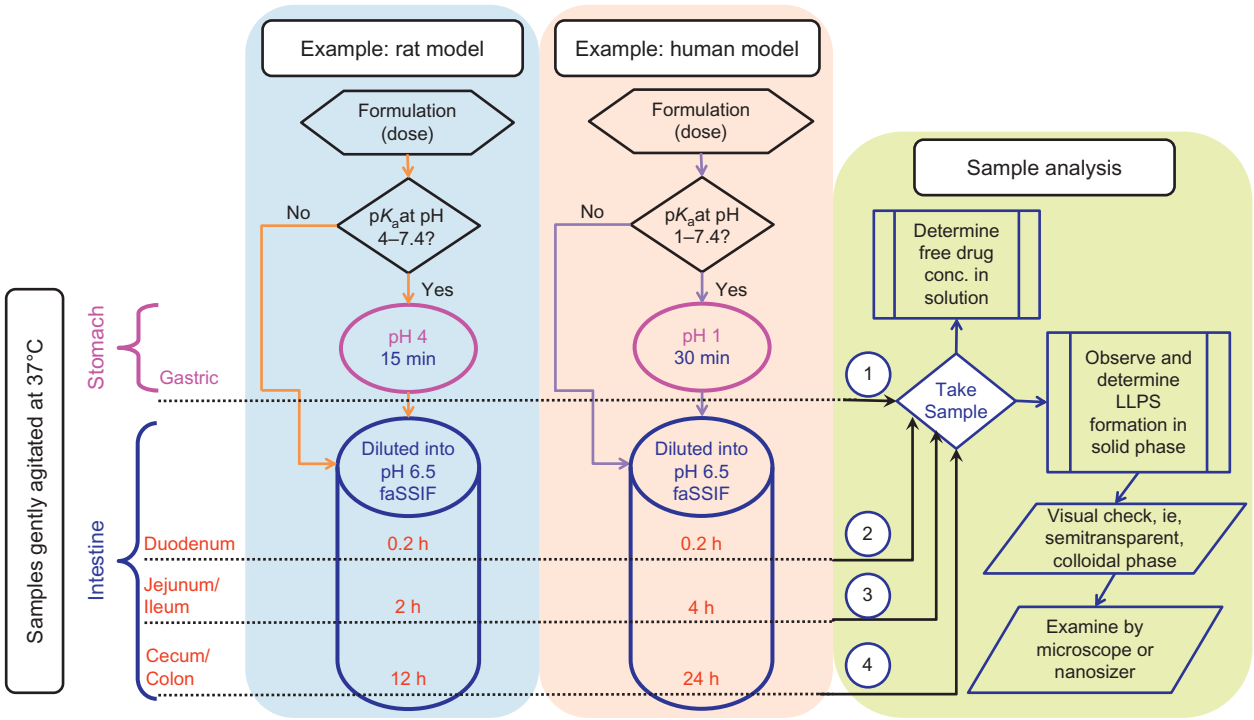


FIGURE 17.11 Schematic illustration of an LLPS-based in vitro dissolution method to evaluate enabling formulations.

illustrated in Table 17.6. If the solubility of the amorphous form in the placebo-containing aqueous medium is the same as that obtained from the dissolution samples, then the amorphous solubility (thus the highest free drug concentration) has been achieved during dissolution. In Table 17.6, the hypothetical formulations #1 and #2 have reached amorphous activity, but #3 did not, although the apparent solution concentration is much higher than those from #1 and #3. In case solubility measurement is also not affordable due to lack of either material or time, as a last resort, the free drug concentration can be inferred based on the residual solid properties. When the solid that is in equilibrium with the solution phase is determined to be amorphous, it can be assumed that the solution concentration is at amorphous solubility. Practically, as long as amorphous droplets or particles are observed, the solution concentration is at amorphous solubility, which is the highest possible supersaturation level. It is more important to know the length of the time the nano or submicron-sized amorphous drug is maintained.

The ranking of the formulation is based on the measurements from both solid-state samples (maintenance of amorphous nanoparticles) and solution-state samples (maintenance of amorphous solubility). For examples, #1 ranked: Nano to submicron-sized amorphous particles are maintained throughout, and the solution concentration reaches amorphous solubility; #2 ranked: Precipitates are in micron size but no crystallization, and the solution concentration reaches amorphous solubility; #3 ranked: Precipitates are initially amorphous but crystallized after 2−3 h into nano or submicron-size crystals, and the solution concentration gradually decreased but not significantly from amorphous solubility; #4 ranked: Massive crystallization into large parcels occurs within 2−3 h, and the solution concentration decreased to significantly below amorphous solubility or close to crystalline solubility.

The LLPS-based in vitro dissolution method is suitable not only for amorphous solid dispersions but also for any other formulations that possibly reach amorphous solubility during dissolution. These formulations include many enabling formulations such as cosolvent/pH adjustment-based formulations, salts, and lipid and surfactant-based formulations.

- Cosolvent/pH adjustment-based formulations. It is easily understandable that these formulations create an extremely high supersaturation when diluted into neutral pH aqueous solutions that can promote LLPS formation if crystallization can be prevented (ie, by adding crystallization inhibitors such polymers).
- Salts. The dissolution of salts has been reported to yield LLPS.[145] Considering that pH_{max} (at which the salts forms) is at least 2 pH units away from pK_a, the apparent solubility of the salt has to be 100-fold higher than the intrinsic solubility of the parent

TABLE 17.6 Author-Created Illustration of Estimating Free Drug Concentration or Solute Activity by Amorphous Solubility Measurement

#	Hypothetical Formulation Composition	Amount Used in Dissolution (mg of Drug)	Solubility Medium (mg of Placebo Dissolved in 5 mL FaSSIF)	Hypothetical Solubility Measured using amorphous form (mg/mL)	Hypothetical Apparent Concentration Measured in Dissolution Samples	Thermodynamic Solute Activity
1	Amorphous solid dispersion: Drug X: HPMCAS = 20: 80 (w/w)	20	80 mg HPMCAS	0.10	0.10	1.0
2	Amorphous solid dispersion: Drug X: HPMCAS: Tween 80 = 20: 65: 15 (w/w)	20	65 mg HPMCAS and 15 mg Tween 80	0.50	0.50	1.0
3	Surfactant formulation: Drug X: PEG400: Tween 20 = 10: 70: 20 (w/w)	20	40 mg Tween 20 and 140 mg PEG 400	2.0	1.0	0.5

This assumes that dissolution is done in 5 mL FaSSF.

crystalline form. The amorphous solubility advantage (solubility ratio of amorphous-to-crystalline form) usually does not exceed 100-fold of crystalline solubility. Thus, salt has a high potential to reach amorphous solubility during dissolution if without crystallization to the parent form. The polymer used as the suspending agent in the salt suspension may just as well act as a crystallization inhibitor to maintain the amorphous solubility.

- Lipid/surfactant-based formulations. The lipid/surfactant-based formulations, especially Type III and IV, contain a cosolvent, which upon dilution highly likely results in a higher supersaturation than the amorphous solubility. Therefore, the LLPS-based dissolution method is an appropriate avenue to evaluate any of these enabling formulations and aid formulation composition optimizations.

In addition, such an LLPS-based dissolution method is universally applicable to all APIs regardless of their permeability. It is a fact that the formulation effort usually does not change the drug permeability. This reality has been often interpreted as "little can be done through formulation if the absorption is limited by permeability," which is now a misnomer after understanding the LLPS and its relationship to drug activity and driving force for permeability. A more appropriate interpretation is that formulation can play a critical role in maximizing the permeability of BCS Class IV compounds in addition to enhancing the solubility and the dissolution rate. As stated repeatedly, the highest thermodynamic solute activity (or the highest free drug concentration), and thus the highest flux (or the highest driving force) for permeability, occurs at amorphous solubility. In the case of poor permeability due to transporters in the gut (ie, efflux), maintaining the highest free concentration (amorphous solubility) is the best way to saturate the enzymes and to enhance drug absorption.

As a conclusion, considerable advances have been made in last 5 years in understanding the dissolution behavior of the enabling formulations. We can expect this trend to continue into the foreseeable future. With new discoveries that never cease, the in vitro dissolution method will keep evolving to more perfection.

17.7 RATIONALE SELECTION OF FORMULATIONS SUITABLE FOR INTENDED STUDIES

The decision on which formulation strategy to choose often resembles a complicated multivariable analysis. This table summarizes the pros and cons of the different formulation approaches, which is helpful in facilitating the decision-making process (Table 17.7).

The timing of a nonclinical study (during discovery or development), dose regiment (single or repeat-dose study), duration of the study, test species, route of administration, and dose range all have a significant impact on the formulation development challenges that a formulator is presented with and can influence the choice of formulation and excipients. Often during the discovery stage, only very small quantities of a test compound are available for formulation development work, and only a very limited amount of time is available to identify an appropriate formulation for an upcoming study. The combination of extremely high doses desired for many safety studies, restrictions on acceptable dose volumes, and very limited drug substance availability oftentimes result in a scenario for a formulator where only a few milligrams of a drug

TABLE 17.7 Comparison Among Different Formulation Approaches.

Formulation Type	Advantages	Potential Issues
Solution (aqueous or aqueous with pH adjustment or cosolvent)	• Maximizes exposure. • Easy to prepare and dose. • Flexibility of combining pH, cosolvency, and complexation to further enhance solubility. • Eliminates exposure variability due to batch-to-batch variation in physical form and particle size.	• Difficult to develop for extremely poorly soluble compounds. • Tolerability issues from solubilization excipients. • Chemical stability issues.
Suspension (neutral solid form or salt)	• Good tolerability. • Supports high dose as not limited by solubility. • Better chemical stability than solution. • Better bridging to clinical solid dosage form (crystalline suspension).	• Exposure impacted by particle size and physical form. • Physical stability issues (form and particle size).
Nanosuspension	• Improved dissolution and enhanced exposure. • Supports high dose. • Good tolerability.	• Physical stability issues. • More complicated development path. • More complicated preparation and scale-up.
Lipid/surfactant-based formulations (forms either emulsion or microemulsion	• Applicable to many poorly soluble compounds. • Improved exposure due to solubilization and potential lymphatic absorption. • Easy to prepare and dose.	• Limited drug load. • Physical instability (adduct formation). • Lack of IVIVR for formulation selection and optimization. • Tolerability when high percentage of surfactants used.
Solid dispersion	• Good tolerability. • Improved exposure by enhancing dissolution and creating high supersaturation. • Broadly applicable with high success rate.	• Potential physical instability. • Limited drug load for compounds with fast crystallization tendency. • Need to develop a suitable suspension for animal dosing. • Scale-up requires development.

substance can be used to test the solubility in multiple test vehicles to achieve the concentration ranges of tens to hundreds of mg/mL. At the same time, the physical properties of the test compounds, such as particle size, crystallinity, and purity, which have a significant impact on the dissolution rate and the solubility in a given vehicle, might not be fully understood and optimized yet and are often not reproducible from batch to batch. Therefore, during the drug discovery stage, a formulator will most likely develop a fit for purpose formulation based on the physical properties of the drug substance at the time and the amount of material and time that are available. The resulting formulation might not be optimized for long-term stability of the drug substance in the formulation or acceptability of the used excipients for repeat dose or long-term studies, for example.

The quantity of the available drug substance and the knowledge and reproducibility of the physical properties increase as a project moves from discovery into development. Once a compound is identified as a candidate for development or even slightly before, a salt and polymorph screen is typically performed to identify a stable crystalline form. This often results in the selection of a crystal form that is different from the crystalline or amorphous phase that was used in the discovery phase, which can result in changed dissolution and solubility properties. As a result, the toxicology formulation may need to be further optimized for GLP safety studies. Depending on how different the crystal form (free base versus salt) or the formulation is, a bridging study might be necessary to evaluate the potential impact on in vivo exposure and the need to adjust the doses for the GLP study.

Other factors to consider during selecting the toxicology dosing formulation are the ease of formulation preparation and the chemical and physical stability of the formulation, which determines the use-time period. These factors are especially important for long-term studies where the dosing formulation has to be prepared many times and should, therefore, be easy to prepare.

There is usually a strong desire for the formulation to be stable for at least 1 week to reduce the frequency of preparation. Scale differences also need to be taken into consideration. A formulation preparation procedure might work well during formulation development on a small, milliliter scale under the employment of certain techniques such as sonication for solubilization or trituration by hand to achieve wetting, particle size reduction, or homogenization. However, it might be very difficult to be reproduced by a CRO (Contract Research Organization) or the drug safety department of the same organization on the larger scale of several liters during the actual toxicology study. Therefore, close communication and alignment between the formulator and the drug safety department/CRO are essential.

Overall, the strategy is to design and develop fit-for-purpose formulations based on the physicochemical properties of the drug candidate, study type, and requirement. The totality of the various important factors is considered to balance quality, speed, and cost when developing a suitable formulation strategy to enable the study.

17.8 CASE STUDY

Herein presented is an example illustrating the process of developing an oral formulation for safety studies.

17.8.1 Model compound properties

Compound X is a molecule with an extremely low aqueous solubility. The pharmaceutical properties are provided in the following table.

Properties	Data
Molecular weight	~500 g/mol
pK_a	8.5 (acidic)
Melting point of crystalline-free acid	~270°C (melt decomposition)
Log*D* at neutral pH	~3
Aqueous solubility	
In aqueous buffers at pH ≤ 8	<12 ng/mL (below detection limit of HPLC assay)
In faSSIF	~0.8 μg/mL
In feSSIF	~5 μg/mL
Chemical stability	Acceptable stability at pH 1–13
	Prone to degradation under peroxide
Permeability by Caco-2	Highly permeable

Based on these properties, Compound X is a BCS II compound. Its poor solubility is due to the combined effect of high lattice energy and hydrophobicity. Enabling technology will be required to deliver this compound. Three formulation approaches were evaluated: salt suspension, lipid/surfactant-based formulations, and amorphous solid dispersion formulations.

17.8.2 Crystallization tendency assessment

Since the three formulation approaches (salt suspension, lipid/surfactant-based formulations, and amorphous solid dispersion formulation) will create a supersaturation upon dosing, the success of these enabling technologies is hinged on how well the crystallization can be inhibited. The effectiveness of four polymers (copovidone, PVP, HPMC E5, and HPMC K3) in prohibiting crystallization was assessed by spiking the concentrated compound stock solution into a pH 6.8 buffer containing predissolved 10 μg/mL of each polymer. The compound concentration was monitored by an UV-dip probe at the ambient temperature. The concentration-time profiles are presented in Fig. 17.12.

The data in Fig. 17.12 indicate that Compound X has a high tendency to crystallize. The concentration in the absence of polymers drops immediately after supersaturation was created, although the concentration did not go all the way to <12 ng/mL (crystalline solubility at pH 6.8), which is attributed to the surface defect of the precipitates that manifest a higher apparent solubility. However, crystallization is significantly inhibited or reduced in the presence of polymers. Any of the four polymers studied was able to maintain

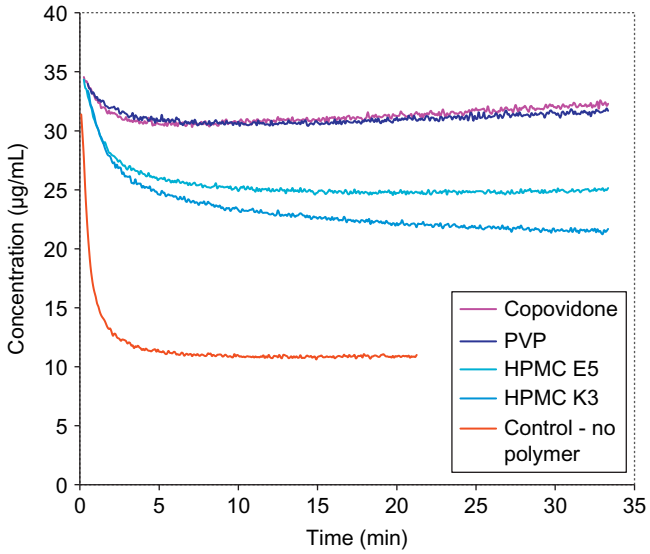


FIGURE 17.12 Polymer screen results in pH 6.8 buffer.

supersaturation. Copovidone and PVP appear to provide a somewhat higher concentration than the HPMC-type polymers, probably due to the nanoparticle effect on the UV-dip probe methodology. Nevertheless, the data suggest that the amorphous solubility in the pH 6.8 buffer is in the range of 22−32 μg/mL, which has been maintained over time by the polymers. Therefore, it is possible to use enabling formulations to deliver this compound.

17.8.3 Development of salt suspension

Taking advantage of the acidic pK_a at 8.5, a salt with acceptable pharmaceutical properties was made successfully. The feasibility of the salt suspension was studied by powder dissolutions in FaSSIF and in a suspending agent (0.2% HPMC in water) at 37°C, with samples filtered and analyzed by HPLC assay. A concentration range of 22−32 μg/mL was maintained for ~30 min in FaSSIF and for several hours in 0.2% HPMC. The 0.2% HPMC suspending agent provided the feasibility to dosing the salt as a suspension.

17.8.4 Development of lipid/surfactant-based formulations

The initial PK study was conducted in the standard PEG 400/DMSO (90/10) vehicle in which the solubility of Compound X was about 5 mg/mL. With the projected dose of ≥60 mg/kg, neat PEG 400 was not able to achieve the required solubility. To develop a viable formulation, the solubility screen was conducted in three groups of excipients: cosolvent, lipid, and surfactant, as shown in Table 17.8.

The solubility data in Table 17.8 pointed out potential ways to formulate a solution formulation: (1) using PEG 400 and pH adjustment, since the addition of NaOH (at 1:1 molar of NaOH:API) to ionize the compound increased the solubility by more than 30-fold in PEG 400; (2) using surfactants, since the aqueous solubility in the presence of 1% Tween or vitamin E TPGS was increased by more than 1000-fold. Vitamin E TPGS has a low miscibility with PEG 400, and it can be replaced by another surfactant, Cremophor RH40.

Based on the miscibility among PEG 400, Tween 20, and Cremophor RH40, two vehicles were prepared and tested for API solubility: (1) PEG 400 and Tween 20 at a 80:20 ratio and (2) PEG 400, Tween 20, and Cremophor RH40 at a 50:20:30 ratio. Adequate solubility of approximately 100 mg/mL was achieved to deliver the required dose. A dilution/dissolution test followed by a PK study in rat showed better in vivo data of PEG 400/Tween 20/Cremophor RH40 than PEG 400/Tween 20. This identified PEG 400/Tween 20/Cremophor RH40 as the lead formulation from the lipid/surfactant-based formulation category.

TABLE 17.8 Solubility Screen in Excipients to Develop Lipid/Surfactant-Based Formulations for Compound X

Excipient	Solubility of Crystalline Form of the Free Acid at Ambient Temperature (mg/mL if not Otherwise Noted)
Ethanol	1−2
PEG 400	~5
PEG 400 with NaOH (1:1 molar ratio with API)	>150
Propylene glycol	~3
A fatty acid	~70 μg/mL
A phospholipid	~8
A medium-chain monoglyceride and diglyceride	<4
1% Tween 20 in pH 6.8 buffer	~25 μg/mL
1% vitamin E TPGS in pH 6.8 buffer	~40 μg/mL
PEG 400/Tween 20 (80/20) with NaOH (1:1 molar ratio with API)	Close to 100 mg/mL
PEG 400/Tween 20/Cremophor RH40 (50/20/30) with NaOH (1:1 molar ratio with API)	~102 mg/mL

17.8.5 Development of amorphous solid dispersions

Fig. 17.12 indicates that polymers can effectively manage the physical stability of amorphous Compound X, which is otherwise a fast crystallizer. Table 17.8 shows that the surfactant can be beneficial in formulating Compound X into amorphous solid dispersions. Taking these two pieces of information together, a prototype formulation with the composition of 25% drug load in copovidone and vitamin E TPGS was designed. The organic solvent screen identified that ethanol, which under moderate heating, was able to dissolve the API as well as the polymer and vitamin E TPGS. The amorphous solid dispersion was made by solvent evaporation and characterized. The results have met all the requirements described in the amorphous characterization scheme in Fig. 17.8. In particular, the following positive results from the two critical attributes have been observed.

1. Physical stability of the ASD suspension. The ASD powders were suspended in a pH 6.8 buffer and gently stirred overnight at the ambient temperature.

The recovered solid remained amorphous, as shown by PXRD (Fig. 17.13).

2. Powder dissolution of ASD. The dissolution was conducted in a pH 6.8 buffer at 37°C with samples filtered and analyzed by HPLC assay. The measured concentration was kept at approximately amorphous solubility, as shown in Fig. 17.14. The concentration measured at 12 h remained at ~22 μg/mL on average.

The results demonstrate that the developed amorphous solid dispersion is suitable for dosing as a suspension in preclinical studies. It was, therefore, selected and advanced to in vivo characterization.

17.8.6 In vivo comparison of different formulations

The three developed enabling formulations (salt suspension, PEG 400/surfactants, and ASD suspension) were compared in several PK studies in rats and dogs at different doses. Fig. 17.15 shows the representative results that were generated in dogs at a 25 mg per dog dose. The three enabling formulations yielded a significantly higher BA than without the formulation effort (free acid suspension). Among them, the solution formulation (PEG 400/Tween 20/Cremophor RH40) is comparable to the amorphous solid dispersion, and, therefore, it was selected as the final formulation from

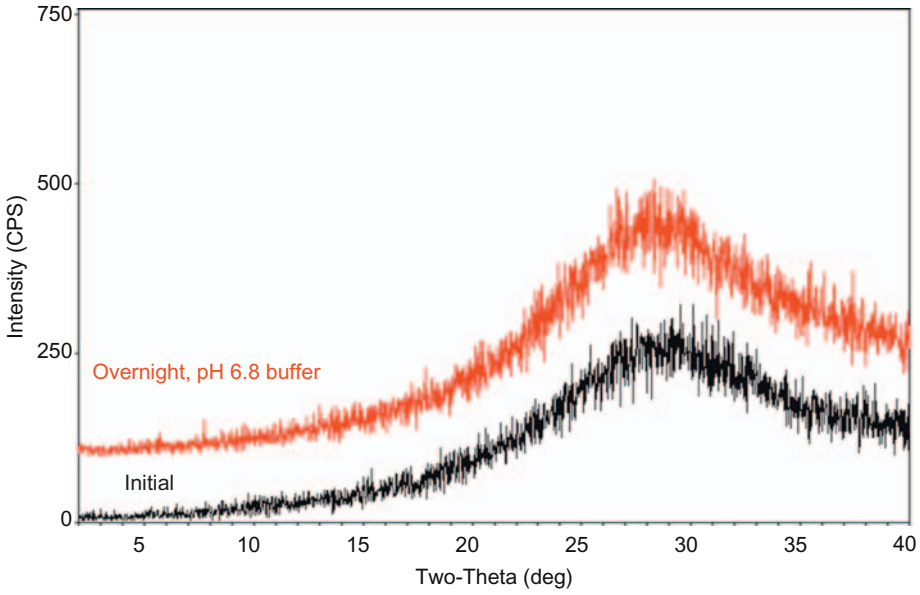


FIGURE 17.13 PXRD pattern of a prototype amorphous solid dispersion of Compound X after suspension in pH 6.8 buffer overnight at ambient temperature.

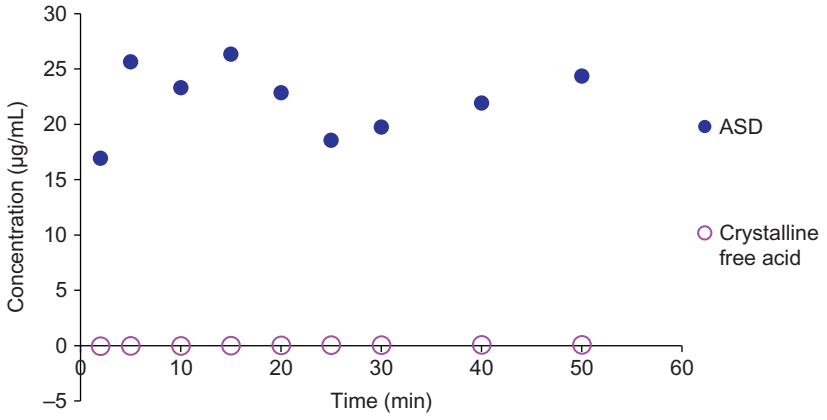


FIGURE 17.14 Powder dissolution of a prototype amorphous solid dispersion of Compound X in pH 6.8 buffer and 37°C. Dissolution profile of the crystalline-free acid was overlaid for comparison.

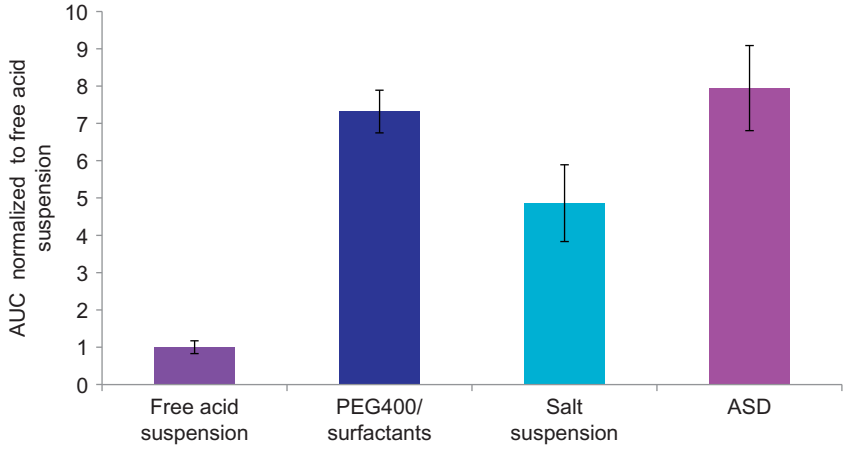


FIGURE 17.15 In vivo comparison among different formulations for Compound X with AUC normalized to that of free acid crystalline form suspension.

the perspective of easier formulation preparation than the ASD. The PEG 400/surfactants formulation was successfully used in long-term GLP studies to evaluate the safety of Compound X.

Acknowledgments

All authors wish to acknowledge the many colleagues and management in their respective companies for the support of their scientific research and publications. In particular, Yi Gao expresses gratitude to Geoff G.Z. Zhang, PhD, at AbbVie for insightful discussions and for sharing the graphic content in Fig. 17.7, and to John Lipari at AbbVie for discussing lipid-based formulations. Christoph Gesenberg would like to thank Kevin Trouba, PhD, DABT; Balvinder Vig, PhD; and Meagan McCallum at Bristol-Myers Squibb for their helpful contributions. Weijia Zheng would like to thank Robert Casella, PhD, at AstraZeneca for valuable discussions.

References

1. Thackaberry EA. Vehicle selection for nonclinical oral safety studies. *Expert Opin Drug Metab Toxicol* 2013;**9**:1635–46.
2. Dressman JB, Yamada K. Animal models for oral drug absorption. In: Welling P, Tse FL, editors. *Pharmaceutical bioeqivalence*. New York, NY: Dekker; 1991. . p. 235–66.
3. Akimoto M, Nagahata N, Furuya A, Fukushima K, Higuchi S, Suwa T. Gastric pH profiles of beagle dogs and their use as an alternative to human testing. *Eur J Pharm Biopharm* 2000;**49**:99–102.
4. Kararli TT. Comparison of the gastrointestinal anatomy, physiology, and biochemistry of humans and commonly used laboratory animals. *Biopharm Drug Dispos* 1995;**16**:351–80.
5. Youngberg CA, Wlodyga J, Schmaltz S, Dressman JB. Radiotelemetric determination of gastrointestinal pH in four healthy beagles. *Am J Vet Res* 1985;**46**:1516–21.
6. Degen LP, Phillips SF. Variability of gastrointestinal transit in healthy women and men. *Gut* 1996;**39**:299–305.
7. Davies B, Morris T. Physiological parameters in laboratory animals and humans. *Pharm Res* 1993;**10**:1093–5.
8. Miyabayashi T. Small intestinal emptying time in normal beagle dogs. A contrast radiology study. *Vet Radiol* 1986;**27**:197.
9. Dressman JB. Comparison of canine and human gastrointestinal physiology. *Pharm Res* 1986;**3**:123–31.
10. Dressman JB, Bass P, Ritschel WA, Friend DR, Rubinstein A, Ziv E. Gastrointestinal parameters that influence oral medications. *J Pharm Sci* 1993;**82**:857–72.
11. Watanabe J, Okabe H, Ichihashi T, Mizojiri K, Yamada H, Yamamoto R. Gastric emptying rate constants after oral administration of drug solution to mice, rats, and rabbits. *Chem Pharm Bull* 1977;**25**:2147–55.
12. Mathers JC, Blake JS. Transit times through the human gut of markers taken at different times of the day. *Proc Nutr Soc* 1983;**42**:111A.
13. <http://www.rivm.nl/interspeciesinfo/intra/rat/stomach/db_rat_stomach.jsp>. Website maintained by National Institute for Public Health and the Environment.
14. Diehl K-H, Hull R, Morton D, Pfister R, Rabemampianina Y, Smith D, et al. A good practice guide to the administration of substances and removal of blood, including routes and volumes. *J Appl Toxicol* 2001;**21**:15–23.
15. Gopinathan S, O'Neill E, Rodriguez LA, Champ R, Phillips M, Nouraldeen A, et al. In vivo toxicology of excipients commonly employed in drug discovery in rats. *J Pharmacol Toxicol Methods* 2013;**68**:284–95.
16. Gad SC, Cassidy CD, Aubert N, Spainhour B, Robbe H. Nonclinical vehicle use in studies by multiple routes in multiple species. *Int J Toxicol* 2006;**25**:499–521.
17. Neervannan S. Preclinical formulations for discovery and toxicology: physicochemical challenges. *Expert Opin Drug Metab Toxicol* 2006;**2**:715–31.
18. Ruble GR, Giardino OZ, Fossceco SL, Cosmatos D, Knapp RJ, Barlow NJ. The effect of commonly used vehicles on canine hematology and clinical chemistry values. *J Am Assoc Lab Anim Sci* 2006;**45**:25–9.
19. Schulze JDR, Peters EE, Vickers AW, Staton JS, Coffin MD, Parsons GE, et al. Excipient effects on gastrointestinal transit and drug absorption in beagle dogs. *Int J Pharm* 2005;**300**:67–75.
20. Anon. Final report on the safety assessment of propylene glycol and polypropylene glycols. *J Am Coll Toxicol* 1994;**13**:437–91.
21. Batrakova EV, Li S, Miller DW, Kabanov AV. Pluronic P85 increases permeability of a broad spectrum of drugs in polarized BBMEC and Caco-2 cell monolayers. *Pharm Res* 1999;**16**:1366–72.
22. Li P, Zhao L. Developing early formulations: practice and perspective. *Int J Pharm* 2007;**341**:1–19.

23. Di L, Fish PV, Mano T. Bridging solubility between drug discovery and development. *Drug Discov Today* 2012;**17**:486–95.
24. Lipinski CA. Drug-like properties and the causes of poor solubility and poor permeability. *J Pharmacol Toxicol Methods* 2000;**44**:235–49.
25. Waring MJ. Lipophilicity in drug discovery. *Exp Opn Drug Discov* 2010;**5**:235–48.
26. Briggner L-E, Hendrickx R, Kloo L, Rosdahl J, Svensson PH. Solid-state perturbation for solubility improvement: a proof of concept. *ChemMedChem* 2011;**6**:60–2.
27. Zornoza A, Martín C, Sánchez M, Vélaz I, Piquer A. Inclusion complexation of glisentide with α-, β- and γ-cyclodextrins. *Int J Pharm* 1998;**169**:239-–244.
28. Hughes JD, Blagg J, Price DA, Bailey S, DeCrescenzo GA, Devraj RV, et al. Physiochemical drug properties associated with in vivo toxicological outcomes. *Bioorg Med Chem Lett* 2008;**18**:4872–5.
29. Pajouhesh H, Lenz GR. Medicinal chemical properties of successful central nervous system drugs. *NeuroRX* 2005;**2**:541–53.
30. Leeson PD, St-Gallay SA. The influence of the 'organizational factor' on compound quality in drug discovery. *Nat Rev Drug Discov* 2011;**10**:749–65.
31. Gleeson MP. Generation of a Set of Simple, Interpretable ADMET Rules of Thumb. *J Med Chem* 2008;**51**:817–34.
32. Lipinski CA. Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov Today Technol* 2004;**1**:337–41.
33. Waring JF, Ciurlionis R, Marsh K, Klein LL, Degoey DA, Randolph JT, et al. Identification of proteasome gene regulation in a rat model for HIV protease inhibitor-induced hyperlipidemia. *Arch Toxicol* 2010;**84**:263–70.
34. Faller B, Ertl P. Computational approaches to determine drug solubility. *Adv Drug Deliv Rev* 2007;**59**:533–45.
35. Hou TJ, Xia K, Zhang W, Xu XJ. ADME evaluation in drug discovery. 4. Prediction of aqueous solubility based on atom contribution approach. *J Chem Inf Comput Sci* 2004;**44**:266–75.
36. Huuskonen J. Estimation of aqueous solubility for a diverse set of organic compounds based on molecular topology. *J Chem Inf Comput Sci* 2000;**40**:773–7.
37. Klopman G, Zhu H. Estimation of the aqueous solubility of organic molecules by the group contribution approach. *J Chem Inf Comput Sci* 2001;**41**:439–45.
38. Ran Y, He Y, Yang G, Johnson JLH, Yalkowsky SH. Estimation of aqueous solubility of organic compounds by using the general solubility equation. *Chemosphere* 2002;**48**:487–509.
39. Wassvik CM, Holmén AG, Bergström CAS, Zamora I, Artursson P. Contribution of solid-state properties to the aqueous solubility of drugs. *Eur J Pharm Sci* 2006;**29**:294–305.
40. Abraham MH, Le J. The correlation and prediction of the solubility of compounds in water using an amended solvation energy relationship. *J Pharm Sci* 1999;**88**:868–80.
41. Delaney JS. ESOL: estimating aqueous solubility directly from molecular structure. *J Chem Inf Comput Sci* 2004;**44**:1000–5.
42. Yan A, Gasteiger J. Prediction of aqueous solubility of organic compounds based on a 3D structure representation. *J Chem Inf Comput Sci* 2003;**43**:429–34.
43. Alsenz J, Kansy M. High throughput solubility measurement in drug discovery and development. *Adv Drug Deliv Rev* 2007;**59**:546–67.
44. Seadeek C, Ando H, Bhattachar SN, Heimbach T, Sonnenberg JL, Blackburn AC. Automated approach to couple solubility with final pH and crystallinity for pharmaceutical discovery compounds. *J Pharm Biomed Anal* 2007;**43**:1660–6.
45. Serajuddin ATM. Salt formation to improve drug solubility. *Adv Drug Delivery Rev* 2007;**59**:603–16.
46. Tong W-QT, Whitesell G. In situ salt screening-A useful technique for discovery support and preformulation studies. *Pharm Dev Technol* 1998;**3**:215–23.
47. Bayomi MA, Abanumay KA, Al-Angary AA. Effect of inclusion complexation with cyclodextrins on photostability of nifedipine in solid state. *Int J Pharm* 2002;**243**:107–17.
48. Challa R, Ahuja A, Ali J, Khar RK. Cyclodextrins in drug delivery: an updated review. *AAPS PharmSciTech* 2005;**6**:E329–57.
49. Amidon GL, Lennernaes H, Shah VP, Crison JR. A theoretical basis for a biopharmaceutic drug classification: the correlation of in vitro drug product dissolution and in vivo bioavailability. *Pharm Res* 1995;**12**:413–20.
50. Food and Drug Administration of United States (2010). Guidance for Industry Waiver of In Vivo Bioavailability and Bioequivalence Studies for Immediate-Release Solid Oral Dosage Forms Based on a Biopharmaceutics Classification System. In *U. S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER)*.
51. Yu LX. An integrated model for determining causes of poor oral drug absorption. *Pharm Res* 1999;**16**:1883–7.
52. Wang J, Douglas RF. Fundamentals of dissolution. In: Qiu Y, Chen Y, Zhang GGZ, editors. Developing solid oral dosage forms pharmaceutical theory and practice *1st edit*. New York, NY: Academic Press; 2009. . p. 309–18.
53. Grant DJW, Higuchi T. Chapter 2 solubility, intermolecular forces, and thermodynamics. In: Saunders WHJ, editor. *Solubility behavior of organic compounds*. New York, NY: John Wiley & Sons; 1990. . p. 12–88.
54. Komasaka T, Fujimura H, Tagawa T, Sugiyama A, Kitano Y. Practical method for preparing nanosuspension formulations for toxicology studies in the discovery stage: formulation optimization and in vitro/in vivo evaluation of nanosized poorly water-soluble compounds. *Chem Pharm Bull* 2014;**62**:1073–82.
55. Gao Y, Olsen KW. Molecular dynamics of drug crystal dissolution: simulation of acetaminophen form I in water. *Mol Pharm* 2013;**10**:905–17.
56. Grant DJW, Brittain HG. Solubility of pharmaceutical solids. *Drugs Pharm Sci* 1995;**70**:321–86.
57. Danhier F, Ucakar B, Vanderhaegen M-L, Brewster ME, Arien T, Preat V. Nanosuspension for the delivery of a poorly soluble anticancer kinase inhibitor. *Eur J Pharm Biopharm* 2014;**88**:252–60.
58. Kumar S, Jog R, Shen J, Zolnik B, Sadrieh N, Burgess DJ. Formulation and performance of danazol nano-crystalline suspensions and spray dried powders. *Pharm Res* 2015;**32**:1694–703.
59. Gould S, Scott RC. 2-Hydroxypropyl-β-cyclodextrin (HP-β-CD): a toxicology review. *Food Chem Toxicol* 2005;**43**:1451–9.
60. Thackaberry EA, Kopytek S, Sherratt P, Trouba K, McIntyre B. Comprehensive investigation of hydroxypropyl methylcellulose, propylene glycol, polysorbate 80, and hydroxypropyl-beta-cyclodextrin for use in general toxicology studies. *Toxicol Sci* 2010;**117**:485–92.
61. Enright BP, McIntyre BS, Thackaberry EA, Treinen KA, Kopytek SJ. Assessment of hydroxypropyl methylcellulose, propylene glycol, polysorbate 80, and hydroxypropyl-β-cyclodextrin for use in developmental and reproductive toxicology studies. *Birth Defects Res, Part B* 2010;**89**:504–16.
62. Bergstrom CAS, Luthman K, Artursson P. Accuracy of calculated pH-dependent aqueous drug solubility. *Eur J Pharm Sci* 2004;**22**:387–98.
63. 8613. Simvastatin. In: O'Neil MJ, Smith A, Heckelman PE, editors. *The merck index an encyclopedia of chemicals, drugs, and biologicals*. 13th ed. Kenilworth, NJ: Merck Research Laboratories; 2001.
64. Alvarez-Lueje A, Valenzuela C, Squella JA, Nunez-Vergara LJ. Stability study of simvastatin under hydrolytic conditions assessed by liquid chromatography. *J AOAC Int* 2005;**88**:1631–6.
65. Ungaro F, Giovino C, Catanzano O, Miro A, Mele A, Quaglia F, et al. Use of cyclodextrins as solubilizing agents for simvastatin: effect of hydroxypropyl-β-cyclodextrin on lactone/hydroxyacid aqueous equilibrium. *Int J Pharm* 2011;**404**:49–56.

66. Pudipeddi P, Serajuddin ATM, Grant DJW, Stahl PH. Solubility and dissolution of weak acids, bases, and salts. In: Stahl PH, Wermuth CG, editors. *Handbook of pharmaceuitcal salts properties, selection, and use*. New York, NY: Wiley-VCH; 2002. p. 19–41.
67. Merritt JM, Viswanath SK, Stephenson GA. Implementing quality by design in pharmaceutical salt selection: a modeling approach to understanding disproportionation. *Pharm Res* 2013;**30**:203–17.
68. Hsieh Y-L, Merritt JM, Yu W, Taylor LS. Salt stability—the effect of phmax on salt to free base conversion. *Pharm Res* 2015;**32**:3110–18.
69. Matsumoto T, Zografi G. Physical properties of solid molecular dispersions of indomethacin with poly(vinylpyrrolidone) and poly(vinylpyrrolidone-co-vinyl-acetate) in relation to indomethacin crystallization. *Pharm Res* 1999;**16**:1722–8.
70. Zhang J, Zografi G. Water vapor absorption into amorphous sucrose-poly(vinyl pyrrolidone) and trehalose-poly(vinyl pyrrolidone) mixtures. *J Pharm Sci* 2001;**90**:1375–85.
71. Alonzo DE, Raina S, Zhou D, Gao Y, Zhang GGZ, Taylor LS. Characterizing the impact of hydroxypropylmethyl cellulose on the growth and nucleation kinetics of felodipine from supersaturated solutions. *Cryst Growth Des* 2012;**12**:1538–47.
72. Liu H, Taylor LS, Edgar KJ. The role of polymers in oral bioavailability enhancement; a review. *Polymer (Guildf)* 2015;**77**:399–415.
73. Cai T, Zhu L, Yu L. Crystallization of organic glasses: effects of polymer additives on bulk and surface crystal growth in amorphous nifedipine. *Pharm Res* 2011;**28**:2458–66.
74. Sun Y, Zhu L, Wu T, Cai T, Gunn EM, Yu L. Stability of amorphous pharmaceutical solids: crystal growth mechanisms and effect of polymer additives. *AAPS J* 2012;**14**:380–8.
75. Powell CT, Cai T, Hasebe M, Gunn EM, Gao P, Zhang G, et al. Low-concentration polymers inhibit and accelerate crystal growth in organic glasses in correlation with segmental mobility. *J Phys Chem B* 2013;**117**:10334–41.
76. Patel DD, Anderson BD. Effect of precipitation inhibitors on indomethacin supersaturation maintenance: mechanisms and modeling. *Mol Pharm* 2014;**11**:1489–99.
77. Chiang P-C, South SA, Daniels JS, Anderson DR, Wene SP, Albin LA, et al. Aqueous versus non-aqueous salt delivery strategies to enhance oral bioavailability of a mitogen-activated protein kinase-activated protein kinase (MK-2) inhibitor in rats. *J Pharm Sci* 2009;**98**:248–56.
78. Chen J, Ormes JD, Higgins JD, Taylor LS. Impact of surfactants on the crystallization of aqueous suspensions of celecoxib amorphous solid dispersion spray dried particles. *Mol Pharm* 2015;**12**:533–41.
79. Higuchi T. Some physical chemical aspects of suspension formulation. *J Am Pharm Assoc (1912-1977)* 1958;**47**:657–60.
80. Ziller KH, Rupprecht H. Control of crystal growth in drug suspensions: design of a control unit and application to acetaminophen suspensions. In: Rubinstein MH, editor. *Pharmaceutical technology: drug stability*. New York, NY: John Wiley and Sons; 1989. p. 118–31.
81. Martin A. Chapter 10 solubility and distribution phenomena. In: Mundorff GH, editor. *Physical pharmacy*. Baltimore, MD: Lippincott Williams & Wilkins; 1993. p. 212–50.
82. Ali HSM, York P, Blagden N, Khoubnasabjafari M, Acree Jr. WE, Jouyban A. Solubility of salbutamol and salbutamol sulphate in ethanol + water mixtures at 25°. *J Mol Liq* 2012;**173**:62–5.
83. Gibson L. Chapter 2 lipis-based excipients for oral drug delivery. In: Hauss DJ, editor. *Oral lipid-based formulations enhancing the bioavailability of poorly water-soluble drugs*. New York, NY: Informa Healthcare USA; 2007. p. 33–61.
84. Strickley RG. Solubilizing excipients in oral and injectable formulations. *Pharm Res* 2004;**21**:201–30.
85. Pouton CW, Porter CJH. Formulation of lipid-based delivery systems for oral administration: materials, methods and strategies. *Adv Drug Delivery Rev* 2008;**60**:625–37.
86. Shah NH, Phuapradit W, Zhang Y-E, Ahmad H, Malick AW. Chapter 6 lipis-based isotropic solutions: design considerations. In: Hauss DJ, editor. *Oral lipid-based formulations enhancing the bioavailability of poorly water-soluble drugs*. New York, NY: Informa Healthcare USA; 2007. p. 129–48.
87. Van den Mooter G. The use of amorphous solid dispersions: a formulation strategy to overcome poor solubility and dissolution rate. *Drug Discov Today Technol* 2012;**9**:e79–85.
88. He Y, Ho C. Amorphous solid dispersions: utilization and challenges in drug discovery and development. *J Pharm Sci* 2015;**104**:3237–58.
89. Yu L. Amorphous pharmaceutical solids: preparation, characterization and stabilization. *Adv Drug Delivery Rev* 2001;**48**:27–42.
90. Brewster ME, Stokbroekx S, Verreck G, Swinney K. Solid dispersions as supersaturating drug delivery systems. *Bull Tech Gattefosse* 2011;**104**:35–50.
91. Newman A, editor. *Pharmaceutical amorphous solid dispersions*. New York, NY: John Wiley & Sons; 2015.
92. Shah N, Sandhu H, Choi DS, Chokshi H, Malick AW, editors. *Amorphous solid dispersions theory and practice*. New York, NY: Springer-Verlag; 2014.
93. Alonzo DE, Zhang GGZ, Zhou D, Gao Y, Taylor LS. Understanding the behavior of amorphous pharmaceutical systems during dissolution. *Pharm Res* 2010;**27**:608–18.
94. Murdande SB, Pikal MJ, Shanker RM, Bogner RH. Solubility advantage of amorphous pharmaceuticals: I. A thermodynamic analysis. *J Pharm Sci* 2010;**99**:1254–64.
95. Alonzo DE, Gao Y, Zhou D, Mo H, Zhang GGZ, Taylor LS. Dissolution and precipitation behavior of amorphous solid dispersions. *J Pharm Sci* 2011;**100**:3316–31.
96. Ilevbare GA, Taylor LS. Liquid-liquid phase separation in highly supersaturated aqueous solutions of poorly water-soluble drugs: implications for solubility enhancing formulations. *Cryst Growth Des* 2013;**13**:1497–509.
97. Raina SA, Van Eerdenbrugh B, Alonzo DE, Mo H, Zhang GGZ, Gao Y, et al. Trends in the precipitation and crystallization behavior of supersaturated aqueous solutions of poorly water-soluble drugs assessed using synchrotron radiation. *J Pharm Sci* 2015;**104**:1981–92.
98. Jackson MJ, Kestur US, Hussain MA, Taylor LS. Dissolution of danazol amorphous solid dispersions: supersaturation and phase behavior as a function of drug loading and polymer type. *Mol Pharm* 2016;**13**:223–31.
99. Zhou D, Zhang GGZ, Law D, Grant DJW, Schmitt EA. Physical stability of amorphous pharmaceuticals: importance of configurational thermodynamic quantities and molecular mobility. *J Pharm Sci* 2002;**91**:1863–72.
100. Gao, Y., Zhang, G. (2015). Oral solid dosage forms comprising a HCV inhibitor, pp. 17pp., Cont.-in-part of U.S. Ser. No. 717,993. AbbVie Inc., USA.
101. Baird JA, Van Eerdenbrugh B, Taylor LS. A classification system to assess the crystallization tendency of organic molecules from undercooled melts. *J Pharm Sci* 2010;**99**:3787–806.
102. Van Eerdenbrugh B, Baird JA, Taylor LS. Crystallization tendency of active pharmaceutical ingredients following rapid solvent evaporation-classification and comparison with crystallization tendency from undercooled melts. *J Pharm Sci* 2010;**99**:3826–38.
103. Ilevbare GA, Liu H, Edgar KJ, Taylor LS. Impact of polymers on crystal growth rate of structurally diverse compounds from aqueous solution. *Mol Pharm* 2013;**10**:2381–93.

104. Gao Y, Olsen KW. Drug-polymer interactions at water-crystal interfaces and implications for crystallization inhibition: molecular dynamics simulations of amphiphilic block copolymer interactions with tolazamide crystals. *J Pharm Sci* 2015;**104**:2132–41.
105. Chen Y, Liu C, Chen Z, Su C, Hageman M, Hussain M, et al. Drug-polymer-water interaction and its implication for the dissolution performance of amorphous solid dispersions. *Mol Pharm* 2015;**12**:576–89.
106. Tian Y, Jones DS, Andrews GP. An investigation into the role of polymeric carriers on crystal growth within amorphous solid dispersion systems. *Mol Pharm* 2015;**12**:1180–92.
107. Dai W-G, Dong LC, Shi X, Nguyen J, Evans J, Xu Y, et al. Evaluation of drug precipitation of solubility-enhancing liquid formulations using milligram quantities of a new molecular entity (NME). *J Pharm Sci* 2007;**96**:2957–69.
108. Yuan X, Xiang T-X, Anderson BD, Munson EJ. Hydrogen bonding interactions in amorphous indomethacin and its amorphous solid dispersions with poly(vinylpyrrolidone) and poly(vinylpyrrolidone-*co*-vinyl acetate) studied using 13C solid-state NMR. *Mol Pharm* 2015;**12**:4518–28.
109. Song Y, Zemlyanov D, Chen X, Nie H, Su Z, Fang K, et al. Acid-base interactions of polystyrene sulfonic acid in amorphous solid dispersions using a combined UV/FTIR/XPS/ssNMR study. *Mol Pharm* 2016;**13**:483–92.
110. Li W, Buckton G. Using DVS-NIR to assess the water sorption behaviour and stability of a griseofulvin/PVP K30 solid dispersion. *Int J Pharm (Amsterdam, Neth.)* 2015;**495**:999–1004.
111. Weuts I, Van Dycke F, Voorspoels J, De Cort S, Stokbroekx S, Leemans R, et al. Physicochemical properties of the amorphous drug, cast films, and spray dried powders to predict formulation probability of success for solid dispersions: Etravirine. *J Pharm Sci* 2011;**100**:260–74.
112. Sotthivirat S, McKelvey C, Moser J, Rege B, Xu W, Zhang D. Development of amorphous solid dispersion formulations of a poorly water-soluble drug, MK-0364. *Int J Pharm (Amsterdam, Neth.)* 2013;**452**:73–81.
113. Qian F, Huang J, Hussain MA. Drug-polymer solubility and miscibility: stability consideration and practical challenges in amorphous solid dispersion development. *J Pharm Sci* 2010;**99**:2941–7.
114. Keratichewanun S, Yoshihashi Y, Sutanthavibul N, Terada K, Chatchawalsaisin J. An investigation of nifedipine miscibility in solid dispersions using raman spectroscopy. *Pharm Res* 2015;**32**:2458–73.
115. Yuan X, Sperger D, Munson EJ. Investigating miscibility and molecular mobility of nifedipine-PVP amorphous solid dispersions using solid-state NMR spectroscopy. *Mol Pharm* 2014;**11**:329–37.
116. Li N, Taylor LS. Nanoscale infrared, thermal, and mechanical characterization of telaprevir-polymer miscibility in amorphous solid dispersions prepared by solvent evaporation. *Mol Pharm* 2016;**13**:1123–36.
117. Sun DD, Lee PI. Haste makes waste: the interplay between dissolution and precipitation of supersaturating formulations. *AAPS J* 2015;**17**:1317–26.
118. Xie T, Taylor LS. Dissolution performance of high drug loading celecoxib amorphous solid dispersions formulated with polymer combinations. *Pharm Res* 2016;**33**:739–50.
119. Xie T, Taylor LS. Improved release of celecoxib from high drug loading amorphous solid dispersions formulated with polyacrylic acid and cellulose derivatives. *Mol Pharm* 2016;**13**:873–84.
120. Knopp MM, Nguyen JH, Mu H, Langguth P, Rades T, Holm R. Influence of copolymer composition on in vitro and in vivo performance of celecoxib-PVP/VA amorphous solid dispersions. *AAPS J* 2016;**18**(2):416–23.
121. Pajula K, Taskinen M, Lehto V-P, Ketolainen J, Korhonen O. Predicting formation and stability of amorphous small molecule binary mixtures from computationally determined flory-huggins interaction parameter and phase diagram. *Mol Pharm* 2010;**7**:795–804.
122. Donnelly C, Tian Y, Potter C, Jones DS, Andrews GP. Probing the effects of experimental conditions on the character of drug-polymer phase diagrams constructed using flory-huggins theory. *Pharm Res* 2015;**32**:167–79.
123. Tao J, Sun Y, Zhang GGZ, Yu L. Solubility of small-molecule crystals in polymers: D-mannitol in PVP, indomethacin in PVP/VA, and nifedipine in PVP/VA. *Pharm Res* 2009;**26**:855–64.
124. Sun Y, Tao J, Zhang GGZ, Yu L. Solubilities of crystalline drugs in polymers: an improved analytical method and comparison of solubilities of indomethacin and nifedipine in PVP, PVP/VA, and PVAc. *J Pharm Sci* 2010;**99**:4023–31.
125. Knopp MM, Olesen NE, Huang Y, Holm R, Rades T. Statistical analysis of a method to predict drug-polymer miscibility. *J Pharm Sci* 2016;**105**(1):362–7.
126. Kyerematen SO, Pudlas M, Woehrle GH. A fast and reliable empirical approach for estimating solubility of crystalline drugs in polymers for hot-melt extrusion formulations. *J Pharm Sci* 2014;**103**:2847–58.
127. Vasconcelos T, Marques S, das Neves J, Sarmento B. Amorphous solid dispersions: rational selection of a manufacturing process. *Adv Drug Delivery Rev* 2016;**100**:85–101.
128. International Conference on Harmonisation. Impurities: guideline for residual solvents. In: *ICH Harmonised Tripartite Guideline Q3c(R5)* 13th edit; 2011.
129. Grodowska K, Parczewski A. Organic solvents in the pharmaceutical industry. *Acta Poloniae Pharm Drug Res* 2010;**67**:3–12.
130. Li N, Ormes JD, Taylor LS. Leaching of lopinavir amorphous solid dispersions in acidic media. *Pharm Res* 2016;**33**(7):1723–35.
131. Singh A, Van den Mooter G. Spray drying formulation of amorphous solid dispersions. *Adv Drug Delivery Rev* 2015;**100**:27–50.
132. Agrawal AM, Dudhedia MS, Zimny E. Hot melt extrusion: development of an amorphous solid dispersion for an insoluble drug from mini-scale to clinical scale. *AAPS PharmSciTech* 2016;**17**:133–47.
133. Gao Y, Carr RA, Spence JK, Wang WW, Turner TM, Lipari JM, et al. A pH-dilution method for estimation of biorelevant drug solubility along the gastrointestinal tract: application to physiologically based pharmacokinetic modeling. *Mol Pharm* 2010;**7**:1516–26.
134. Mathias NR, Xu Y, Patel D, Grass M, Caldwell B, Jager C, et al. Assessing the risk of pH-dependent absorption for new molecular entities: a novel in vitro dissolution test, physicochemical analysis, and risk assessment strategy. *Mol Pharm* 2013;**10**:4063–73.
135. Miller JM, Beig A, Krieg BJ, Carr RA, Borchardt TB, Amidon GE, et al. The solubility-permeability interplay: mechanistic modeling and predictive application of the impact of micellar solubilization on intestinal permeation. *Mol Pharm* 2011;**8**:1848–56.
136. Miller JM, Dahan A. Predicting the solubility-permeability interplay when using cyclodextrins in solubility-enabling formulations: model validation. *Int J Pharm (Amsterdam, Neth.)* 2012;**430**:388–91.
137. Miller JM, Beig A, Carr RA, Webster GK, Dahan A. The solubility-permeability interplay when using cosolvents for solubilization: revising the way we use solubility-enabling formulations. *Mol Pharm* 2012;**9**:581–90.
138. Grundy JS, Anderson KE, Rogers JA, Foster RT. Studies on dissolution testing of the nifedipine gastrointestinal therapeutic system. I. Description of a two-phase in vitro dissolution test. *J Control Release* 1997;**48**:1–8.
139. Shi Y, Gao P, Gong Y, Ping H. Application of a biphasic test for characterization of in vitro drug release of immediate release formulations of celecoxib and its relevance to in vivo absorption. *Mol Pharm* 2010;**7**:1458–65.

140. Mudie DM, Shi Y, Ping H, Gao P, Amidon GL, Amidon GE. Mechanistic analysis of solute transport in an in vitro physiological two-phase dissolution apparatus. *Biopharm Drug Dispos* 2012;**33**:378–402.
141. Augustijns P, Brewster ME. Supersaturating drug delivery systems: fast is not necessarily good enough. *J Pharm Sci* 2012;**101**:7–9.
142. Raina SA, Zhang GGZ, Alonzo DE, Wu J, Zhu D, Catron ND, et al. Enhancements and limits in drug membrane transport using supersaturated solutions of poorly water-soluble drugs. *J Pharm Sci* 2014;**103**:2736–48.
143. Raina SA, Zhang GGZ, Alonzo DE, Wu J, Zhu D, Catron ND, et al. Impact of solubilizing additives on supersaturation and membrane transport of drugs. *Pharm Res* 2015;**32**:3350–64.
144. Taylor LS, Zhang GGGZ. Physical chemistry of supersaturated solutions and implications for oral absorption. *Adv Drug Delivery Rev* 2016;**101**:122–42.
145. Almeida e Sousa L, Reutzel-Edens SM, Stephenson GA, Taylor LS. Supersaturation potential of salt, co-crystal, and amorphous forms of a model weak base. *Cryst Growth Des* 2016;**16**:737–48.

This page intentionally left blank

CHAPTER

18

Rational Design for Amorphous Solid Dispersions

A. Newman

Seventh Street Development Group, Kure Beach, NC, United States

18.1 INTRODUCTION

Most active pharmaceutical ingredients (APIs) are crystalline materials that possess long-range order and well-defined structures. Many properties, such as solubility, stability, and bioavailability (BA) can be related back to the crystalline form.[1] In contrast, amorphous materials do not possess the long-range order of crystals, but will exhibit short-range order (Fig. 18.1).[2]

Amorphous solids are commonly used for improving dissolution and oral absorption of poorly soluble compounds because they exhibit higher apparent solubility values and faster dissolution rates compared with crystalline materials. However, they also have disadvantages, such as poor physical stability, which can result in crystallization, and poor chemical stability, which can lead to degradation. Since they are metastable forms, development of amorphous compounds involves a greater understanding of the handling, storage, and processing conditions needed to prevent crystallization and maintain the amorphous state.

Amorphous solid dispersions are defined as an amorphous API mixed with a polymer or mixtures of polymers and surfactants.[3,4] The polymer stabilizes the amorphous drug by lowering the chemical potential of the drug and increasing the activation energy needed for crystallization (Fig. 18.2),[5] which results in better physical and chemical stability, higher apparent solubility, and faster dissolution. Dispersions can also display extended supersaturation in solution due to crystallization inhibition by the polymers. By improving the stability and apparent solubility, amorphous solid dispersions can be more easily processed into common solid oral dosage forms. Due to the low equilibrium solubility of many APIs in the polymer,[6] and the amorphous nature of the API in the dispersion, these materials are generally in a metastable state and, therefore, may still exhibit poor physical and chemical stability.

This chapter introduces amorphous solid dispersions and their applications in designing solid oral dosage forms during drug development. It covers the key components in a dispersion, physical characterization techniques used to analyze the materials, as well as screening and selection methodology to help find a dispersion with the best properties, such as physical stability and solubility. It also includes a brief introduction of common processing technologies for manufacturing amorphous solid dispersions and examples of marketed products. More in-depth information on the fundamental principles, materials, formulation and process development, characterization, and regulatory considerations in developing amorphous solid dispersions are can be found in chapters "Crystalline and Amorphous Solids," "Solid State Characterization and Techniques," and "Oral Formulations for Preclinical Studies: Principle, Design and Development Considerations" of this book, recent books, and review articles focused on amorphous solid dispersions.

18.2 KEY COMPONENTS OF AMORPHOUS SOLID DISPERSIONS

The most important components of a stable and effective amorphous solid dispersion are carrier polymer(s) and surfactant(s). The choice and concentration of the polymer can greatly influence the stability and properties

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00018-2

of the dispersion. Early reports referred to solid dispersions as mixtures of polymer and crystalline drug.[7] The small particle size of the crystalline drug would help improve the dissolution for many systems. The term *amorphous solid dispersion* has been used to describe solid mixtures of polymer with amorphous drug. A number of other terms have been used to describe these systems, such as solid dispersion,[8] solid solution,[9] and molecular dispersions.[10] It is important to determine if the material being described in a study contains crystalline or amorphous API since the nomenclature has not been standardized in the literature.

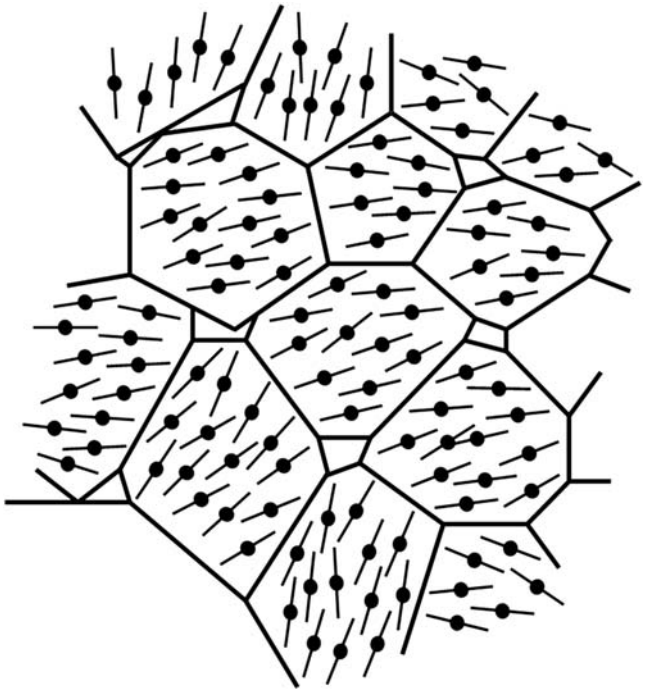


FIGURE 18.1 Short-range order in amorphous material. *Source: Reproduced with permission from Bates S, Zografi G, Engers D, Morris K, Crowley K, Newman A. Analysis of amorphous and nanocrystalline solids from their X-ray diffraction patterns.* Pharm Res *2006;23(10):2333–49.*

Polymers are the most common excipients incorporated into amorphous solid dispersions and a wide variety of polymers are available.[11] Other excipients, such as cyclodextrins[12] and phospholipids,[13] have been added to dispersion systems; it should be noted that both of these excipients provide a different solubilization mechanism than polymers (Fig. 18.3). Small molecules, such as citric acid[15] and amino acids,[16] have also been used to produce amorphous solid dispersions. The amorphous components in the dispersion should be miscible, meaning that the drug and polymer are composed of a single chemically homogeneous phase where all of the components are intimately mixed at the molecular level and the properties of the blend are different than the properties of the pure components. In contrast, a physical mixture of amorphous API and polymer would have the same properties as the pure components. Many miscible amorphous dispersions have improved physical stability,[17] but can become immiscible under certain conditions, such as elevated relative humidity where water is absorbed to produce a ternary system.[18] Partial miscibility is also observed and can be dependent on the concentrations used. For example, felodipine was found to be partially miscible with polyacrylic acid[19] and the dispersion was found to readily crystallize.[17] In contrast, felodipine was found to be miscible with polyvinylpyrrolidone (PVP) and this polymer proved to be an effective crystallization inhibitor.[20] Miscibility can be investigated using a variety of characterization techniques, such as differential scanning calorimetry (DSC) and spectroscopic methods.[21] Amorphous solid dispersions have been utilized at various points in the development process. For early animal studies, dispersions can provide improved apparent solubility resulting in a wide range of concentrations required for toxicology and BA work. Suspensions of amorphous dispersions have also been used for BA studies; it should be noted that the propensity of the dispersion

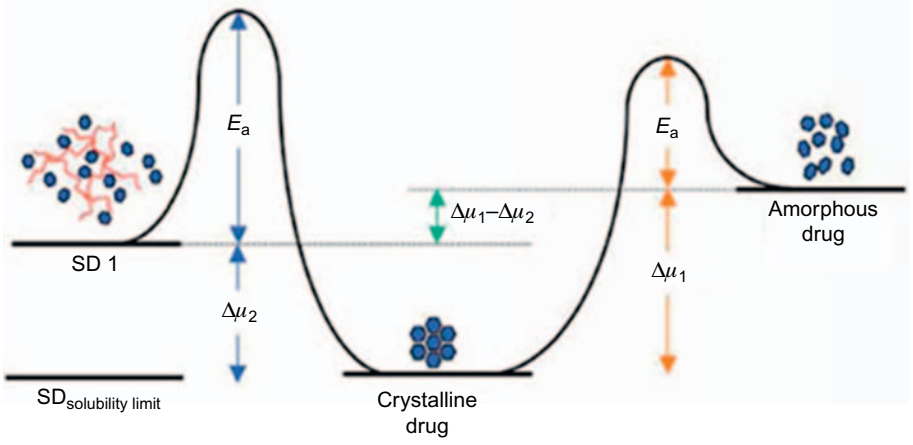


FIGURE 18.2 Hypothetical energy temperature diagram for amorphous, crystalline, and amorphous solid dispersions.[5]

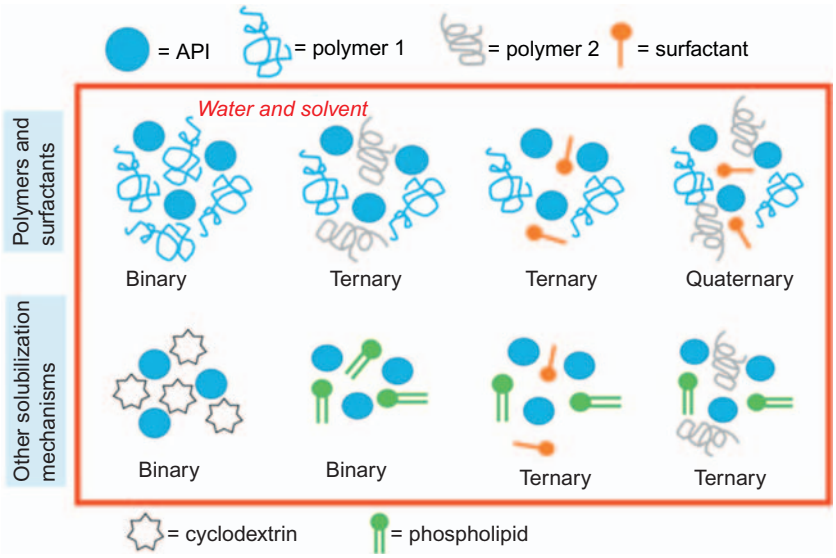


FIGURE 18.3 Graphic of different types of amorphous solid dispersions. The red (gray in print versions) box indicates that properties can change based on solvent and water content.[14]

to crystallize and the amount of API dissolved in the carrier should be understood for these preparations. If the dispersions provide the needed improvements in the preclinical studies, they can be moved into early clinical trials with simple formulations. Suspensions and neat dispersion in a capsule are formulations that have been used for early human trials. In some cases, excipients such as disintegrants or flow agents are added to provide a slightly more complex formulation with better properties than the dispersion alone. The amount of formulation development depends on the characteristics of the dispersion and the needs of the study. For some systems, the amorphous dispersion may have liabilities that do not allow the production of a marketed product. For these cases, a viable solid form is needed and additional work around a different form and/or formulation may be needed. There are a number of cases where amorphous solid dispersions have been used in late clinical trials and developed into marketed products, as shown in Table 18.1. Some dispersions may need special handling, such as using anhydrous excipients to keep water levels low, while others may be more robust, allowing for a variety of excipients and formulation processes.

18.3 CHARACTERIZATION OF AMORPHOUS DISPERSIONS

Many of the characterization methods used for solid form analysis can be used for amorphous solid dispersions. This section will cover common methods such as X-ray powder diffraction (XRPD), DSC, and spectroscopy infrared (IR), Raman, and solid-state nuclear magnetic resonance (SSNMR). Details on these methods can be found in chapter "Solid State Characterization and Techniques"; information provided here will be specific to amorphous solid dispersions. A number of emerging analytical methods have also been discussed.[22]

XRPD of amorphous dispersions should result in an amorphous halo if both the drug and polymer are amorphous. Crystalline peaks from the drug should not be evident in freshly prepared samples. XRPD can be used to monitor dispersion samples over time to determine conditions where the amorphous dispersion is stable (remains amorphous) or unstable (drug crystallizes). A 1:4 nifedipine:PVP amorphous solid dispersion compressed into tablets and stored at 60°C/75% relative humidity (RH) was analyzed by XRPD and dissolution over a period of 2 weeks.[23] Crystalline material was evident from XRPD after 1 week and was correlated with a significant slowdown in dissolution when compared with the initial amorphous dispersion.

The amorphous halo observed in the XRPD pattern can help determine miscibility. A physical mixture of the amorphous drug and polymer will be readily described by the individual components. When the components are miscible and interact with each other, the scattering pattern is different and will produce a different halo. Computational methods, such as pair distribution functions[19,24] and the pure curve resolution method,[25] can also be used with the XRPD data to help determine miscibility of a dispersion.

DSC and modulated DSC (mDSC) are commonly used for characterization of dispersions. Small amounts of crystalline material can be readily observed from a melt endotherm of the crystalline solid.[26] When the drug is fully amorphous, the DSC will show two glass transition

TABLE 18.1 Marketed Products Using Amorphous Solid Dispersions[14]

Product	API	Company	BCS class	Polymer	Dispersion process
Afeditab	Nifedipine	Elan/Watson	2	Poloxamer or PVP	Melt/absorb on carrier
Certican (non-US)	Everolimus	Novartis	3	HPMC	Melt or spray drying
Cesamet	Nabilone	Valeant Pharmaceuticals	2 or 4	PVP	Melt extrusion
Fenoglide	Fenofibrate	LifeCycle Pharma	2	PEG	Spray melt
Gris-PEG	Griseofulvin	Novartis/Pedinol	2	PEG	Melt extrusion
Ibuprofen	Ibuprofen	Soliqs	2	various	Melt extrusion
Incivek (US)	Telaprevir	Vertex Pharmaceuticals	2 or 4	HPMCAS	Spray drying
Incivo (Europe)	Telaprevir	Janssen Pharmaceuticals	2 or 4	HPMCAS	Spray drying
Intelence	Etravirine	Janssen Pharmaceuticals	4	HPMC	Spray drying
Isoptin SRE-240	Verapamil	AbbVie Inc	2	HPC/HPMC	Melt extrusion
Kalydeco	Ivacaftor	Vertex	2 or 4	HPMCAS	Spray drying
Kaletra	Lopinavir and ritonavir	AbbVie	2 and 4	PVP VA	Melt extrusion
LCP-Tacro	Tacrolimus	LifeCycle Pharma/Veloxis	2	HPMC	Melt granulation
Nimotop	Nimodipine	Bayer	2	PEG	Spray drying/fluid bed
Nivadil	Nilvadipine	Fujisawa Pharmaceutical	2	HPMC	
Norvir	Ritonavir	AbbVie	4	PVP VA	Melt extrusion
Noxafil	Posaconazole	Merck	2	HPMCAS	Melt extrusion
Onmel	Itraconazole	GlaxoSmithKline/Stiefel	2	PVP VA	Melt extrusion
Prograf	Tacrolimus	Astellas Pharma Inc	2	HPMC	Spray drying/fluid bed
Rezulin[a]	Troglitazone	Pfizer (Parke-Davis)	2	PVP	Melt extrusion
Sporonox	Itraconazole	Janssen Pharmaceuticals	2	HPMC	Spray layering (fluid granulation)
Torcetrapib[b]	Torcetrapib	Pfizer	2	HPMCAS	Spray drying
Zelboraf	Vemurafenib	Roche	4	HPMCAS	Antisolvent precipitation
Zortress (US)	Everolimus	Novartis Pharmaceuticals	3	HPMC	Melt or spray drying

[a]*Recalled in 2000 due to toxicity issues.*
[b]*Development halted during Phase III clinical trials.*

temperatures (T_gs) for an immiscible system and one T_g for a miscible system.[21] The T_g of the miscible system will be altered depending on the API:polymer concentration and can be estimated using the Fox equation.[27] In some cases, such as the trehalose-dextran system, one T_g will be observed for an immiscible system due to the limitations of the DSC technique based on domain size.[24] Immiscibility was confirmed for the trehalose-dextran system using pair distribution functions[24] and SSNMR.[28]

It is important to note that water or solvent content can also change the T_g of a sample. A rule of thumb is that 1% water can lower a T_g value by approximately 10°C,[29] and the Fox equation can be used to estimate the T_g of a binary dispersion (API and polymer) containing water or solvent. Care should be taken to measure the water/solvent content of the sample before collecting DSC data so the T_g value can be correlated to the solvent content. Sample preparation is also key to getting reproducible T_g measurements. For a "wet" T_g (containing water or solvent), the sample should be placed in a hermetically sealed pan without a pinhole for the DSC scan; this preparation keeps the volatiles in the sample pan throughout the run.[30] For a "dry" T_g, temperature cycling is used to first remove the volatiles with subsequent T_g measurement for the solvent-free sample.[31]

IR and Raman spectroscopy are commonly used to probe intermolecular interactions between the components of a dispersion. The different selection rules for these techniques often result in the API having a stronger Raman response than the polymer, while polymers and drugs will typically both have strong IR absorption peaks that result in more signal overlap for some systems.

A number of studies have used IR spectroscopy to determine interactions and miscibility in dispersions.

This technique was used for an indomethacin-PVP system to show that dimers present in the crystalline drug were not present in the dispersion, but, were replaced by intermolecular bonding between the indomethacin and the polymer.[32] The hydrogen bonding observed in the dispersion helped inhibit crystallization of the drug by preventing indomethacin dimer formation. Other studies have utilized transmission IR spectroscopy of thin dispersion films of droperidol, ketoprofen, nifedipine, and pimozide with PVP, with all showing complete miscibility of the components.[33] Upon exposure to elevated relative humidity, it was found that the droperidol, nifedipine, and pimozide dispersions became immiscible forming drug-rich phases that eventually crystallized. Another study investigated the hydrogen bonding of felodipine with different polymers (PVP, hydroxypropyl cellulose (HPMC), and hydroxypropyl methylcellulose acetate succinate (HPMCAS)).[20] Based on the peak shifts, it was found that PVP formed the strongest hydrogen bonds with felodipine.

The use of Raman spectroscopy to analyze dispersion samples is also growing. As with IR spectroscopy, it is used to look at interactions between components in dispersions and determine miscibility. One report investigating diflunisal in dispersions containing PVP, HPMCAS, and HPC showed phase separation in the HPMCAS dispersions only.[34] Physical stability of dispersions has been investigated with Raman spectroscopy by quantifying the amount of crystalline material present. An example is an ibininabant:PVP dispersion formulated into a tablet with a 1% drug loading[35] where it was possible to detect the presence of 5% crystalline material (less than 0.05% of the total sample mass). While this result is not typical for many Raman methods, it does illustrate the possibilities when there is strong scattering of the drug and little overlap with the polymer. Raman microscopy has also differentiated amorphous and crystalline domains of a dispersion, as reported for a troglitazone:PVP system.[36]

Solid-state NMR has been used to investigate a variety of amorphous solid dispersion properties, such as miscibility,[28,37] intermolecular interactions,[38] molecular mobility,[39] domain size,[28] crystallization kinetics,[40] and ionization.[41] The technique has also been used to identify[28] and quantify crystalline material[42] in dispersion samples.

While ^{13}C SSNMR data are commonly collected for pharmaceutical compounds, a variety of other atoms can also be used to gain information on dispersions, such as ^{19}F[42] and ^{15}N.[41] ^{15}N atoms exhibit a low natural abundance, as does ^{13}C, so larger rotor sizes, more sample, and longer acquisition times are needed. ^{19}F atoms exhibit high sensitivity and high selectivity in fluorinated drugs and SS19F NMR analyses are becoming more common as the number of drugs containing fluorine continues to rise. One study used ^{19}F to investigate the mobility of flufenamic acid at a 20% drug loading in PVP or HPMC dispersions.[40] It was found that the molecular mobility of the flufenamic acid was higher in the HPMC dispersion and was related back to lower physical stability (enhanced crystallization) compared with the PVP dispersion. Another study discusses the quantitation of crystalline material in an AMG517:HPMCAS dispersion using ^{19}F SSNMR.[42] A detection limit of 3% crystalline material was measured for the dispersion at a variety of drug loadings.

Collecting the moisture sorption/desorption curves for amorphous solid dispersions provides useful information on hygroscopicity, handling conditions, and possible phase transformations. Data are commonly collected using an automated moisture sorption system or static humidity chambers that employ salt solutions.[43] The type of sorption/desorption obtained can then be used to probe the properties of the dispersion. In a study with various drugs and the hygroscopic polymer PVP, it was found that some miscible dispersions exposed to elevated relative humidity resulted in crystalline drug being produced (ketoprofen and indomethacin dispersions) while pimozide:PVP dispersion separated into a physical mixture of amorphous drug and amorphous polymer[33] A follow-up study with less hygroscopic polymers found that systems with strong drug-polymer interactions and a less hygroscopic polymer (such as HPMCAS) are less susceptible to moisture-induced phase separation, while more hydrophobic drugs are more susceptible even at low water contents.[44] In addition, crystallization kinetics have been determined using water uptake data.[45]

18.4 SCREENING AND SELECTION OF AMORPHOUS SOLID DISPERSIONS

Amorphous solid dispersion screens are used to find possible API and polymer combinations that provide the necessary properties for development. There are a number of steps in the screening and selection process, as outlined in Fig. 18.4. For an amorphous dispersion screen, collecting information on the API is the first step and data may include chemical structure, solid form, physical stability, chemical stability, hydrogen bond donors and acceptors, calculated solubility parameters, log P, and known solubilities of the solid in solvents and biorelevant media. All of this information may not be available for early development compounds and may be determined during the screen or later in development. Characterization of the API will provide a baseline for comparison when dispersions are found and can be used to help design experiments, such as the melting point of the API for melt

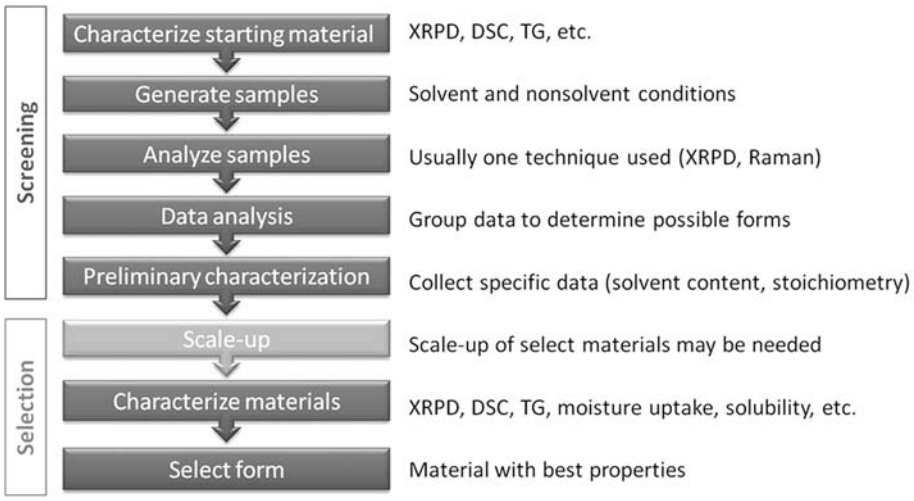


FIGURE 18.4 Outline of screening and selection activities.[46]

experiments and the temperature of desolvation or decomposition for heating studies. It is also useful to prepare the amorphous API for comparison with any dispersions that may be produced.

Designing the screening experiments will include a number of steps, such as choosing excipients, API: polymer ratios, and experimental conditions. Polymer properties to consider may include molecular weight, T_g, hydrophilic/hydrophobic properties, hydrogen bond donors/acceptors, calculated solubility parameter, pH-solubility profile, thermal stability, water uptake, and monomer levels. A variety of methods have been reported for choosing polymers to be included in a screen and some of these are summarized in Table 18.2. Some of these approaches can be easily incorporated into most screens and some will require more time and expertise. It is important to understand the API properties that need to be changed and try to include polymers that may improve that property. For example, if an API has a low T_g then a polymer with a significantly higher value would be used to help increase the T_g of the dispersion, or if the API is hygroscopic then a nonhygroscopic polymer would be targeted. For early development compounds, binary mixtures of API and polymer will be the easiest to investigate, but the option of adding multiple polymers or other components, such as surfactants, to make ternary systems can also be considered. Adding ternary systems to the screen will increase the number of samples and add complexity, but should be considered for difficult systems. The ratio of API:polymer also needs to be established and many studies use a range of 50–70% polymer, but it has been reported that even small amounts of polymer (such as 5%) can provide stabilization for up to 6 months and 30% polymer can provide stability for at least 2 years.[52] A range of compositions should be explored since it has been found that miscibility may change with API:excipient ratio.[53] The estimated drug loading should also be considered because the amount of polymer can significantly increase the size of the capsule or tablet to be produced.

Screening methods typically involve melt or solvent techniques, and many screening methods can be performed in 96-well plates.[54,55] To use melting methods, it needs to be established that the API will not thermally decompose at elevated temperature. Thermal screening methods include melt mixing,[56] DSC,[57] melt compression,[54] and small-scale extruders.[57] For solvent-based methods, it is important to find a solvent or solvent mixture where all components are soluble and chemically stable. When designing screening experiments, the goal is to remove the solvent as quickly as possible to produce the metastable amorphous dispersion. Methods include rotary evaporation,[47] freeze drying,[47] film preparation,[54] electrospinning,[58] and spray drying.[48] In many screens, a tiered approach is used and a combination of methods may be used to make different amounts of material throughout the screening and selection process. For example, a combination of DSC (5–10 mg), melt mixing (10 g), and small-scale extrusion (>10 g) was used to explore dispersions of PVP and five pharmaceutical compounds.[57]

Initial testing of the dispersions begins with XRPD to determine which samples are amorphous. From there a variety of properties can be evaluated, such as T_g, solvent content, solubility, hygroscopicity, dissolution, and stability. The amount of initial testing may be limited by the amount of material available, so it is important to understand the desired properties of the dispersion and prioritize these as the initial set of tests when possible. Based on these properties, a smaller set of amorphous solid dispersions will be moved to the next stage, involving scale-up and additional testing.

TABLE 18.2 Approaches for Choosing Polymers in a Dispersion Screen[14]

Method	Description	Comments	Reference
Empirical	Choose common polymers	Use a variety of polymers; melting point or solubility may be considered	47
Manufacturing	Choose polymers based on the desired large scale manufacturing (melt extrusion and spray drying)	Need low-melting polymers for melt extrusion, need solubility in solvent for spray drying	48,49
Solubility parameter	Choose polymers with a similar solubility parameter as the API	Similar solubility parameter maximizes mixing and promotes miscibility	a,b
Miscibility	Choose polymers that are miscible with API	Miscible systems may show better physical stability	c
Polymer properties	Choose polymer with high T_g	High T_g of polymer will increase T_g of the dispersion; want T_g around 70–80°C to be about 50°C above RT	21,47
	Choose polymer with low water sorption	Water will lower T_g and may decrease physical stability	47, d
	Review pH where polymer dissolves compared to pH of targeted absorption site in GI tract	Compare pH with that of API properties (acid vs base) and site of absorption (gastric vs intestinal)	50, e
DSC	Determine T_m/T_g ratio	Choose polymers with high T_g values to increase T_g of the dispersion compared to the amorphous API:high ratios may crystallize more easily	f
	Reduced crystallization temperature	Crystallization peak in DSC increases in temperature with increased interactions between API and polymer	a
	Determine width of T_g	Wide T_g may indicate partial miscibility	g
Interactions	Look at common H-bonding motifs, ion dipole interactions, or acid-base interactions between drug and polymer	Try to maximize interactions in dispersion disrupt bonding in crystalline materials (eg, PVP disrupts indomethacin dimers)	32, d,h,i
	Acid-base interactions	Ionic interactions produce strong interactions for stable amorphous solid dispersions	i,j
	Complex interactions	Use GEMANOVA analysis based on XRPD microscopy data	k
	Ab initio calculations	Rank order crystallization inhibitors	h
	Determine binding constants using affinity capillary electrophoresis	Drug-polymer binding constants based on changes in effective electrophoretic mobility of the drug upon binding to the polymer	l
Supersaturation	Determine polymers that improve supersaturation or stability in solution	Stable supersaturated solutions provide API:polymer ratios; spray drying used to produce dispersions	51, m,n
Drug solubility in polymer	Determine level of drug that will dissolve in polymer	Gives upper concentration for unsaturated dispersion	6, c,o

[a] *Bhugra C, Schwabe R, Telang C, Li Z. Rational selection of polymers for amorphous dispersions: using reduced crystallization temperature to assess stability potential. 2012 AAPS National Meeting Poster Presentation, Chicago, IL.*

[b] *Shah N, Sandhu H, Choi DS, Kalb O, Page S, Wyttenbach N. Chapter 8. Structured development approach for amorphous systems. In: Williams III RO, Watts AB, Miler DA, editors.* Formulating poorly water soluble drugs. *New York: Springer; 2012. p. 267–310.*

[c] *Marsac PJ, Shamblin SL, Taylor LS. Theoretical and practical approaches for prediction of drug-polymer miscibility and solubility.* Pharm Res *2006;23(10):2417–26.*

[d] *Kaushal AM, Gupta P, Bansal AK. Amorphous drug delivery systems; molecular aspects, design, and performance.* Crit Rev Therap Drug Carrier Sys *2004;21(3):133–93.*

[e] *Curatolo W, Nightingale JA, Herbig SM. Utility of Hydroxypropylmethylcellulose Acetate Succinate (HPMCAS) for initiation and maintenance of drug supersaturation in the GI milieu.* Pharm Res *2009;26:1419–31.*

[f] *Friesen DT, Shanker R, Crew M, Smithey DT, Curatolo WJ, Nightingale JAS. Hydroxypropyl methylcellulose acetate succinate-based spray-dried dispersions: an overview.* Mol Pharm *2008;5(6):1003–19.*

[g] *Song M, Hammiche, A, Pollock HM, Hourston DJ, Reading M. Modulated differential scanning calorimetry: 4. miscibility and glass transition behaviour in poly(methyl methacrylate) and poly(epichlorohydrin) blends. Polymer 1996;37(25):5661–5.*

[h] *Van Eerdenbrugh B, Taylor LS. An ab initio polymer selection methodology to prevent crystallization in amorphous solid dispersions by application of crystal engineering principles. CrystEngComm 2011;13:6171–8.*

[i] *Yoo S-U, Krill SL, Wang, Telang C. Miscibility/stability considerations in binary solid dispersion systems composed of functional excipients towards the design of multi-component amorphous systems.* J Pharm Sci *2009;98(12):4711–23.*

[j] *Van Eerdenbrugh B, Taylor LS. Small scale screening to determine the ability of different polymers to inhibit drug crystallization upon rapid solvent evaporation.* Mol Pharm *2010;7 (4):1328–37.*

[k] *Wu JX, Van Den Berg F, Sogaard SV, Rantanen J. Fast-track to a solid dispersion formulation using multi-way analysis of complex interactions.* J Pharm Sci *2013;102(3):904–14.*

[l] *Jia J, Choi DS, Chokshi H. Determination of drug-polymer binding constants by affinity capillary electrophoresis for aryl propionic acid derivatives and related compounds.* J Pharm Sci *2013;102(3):960–6.*

[m] *Janssens S, Nagels S, Novoa de Armas H, D'Autry W, Schepdael A, Van den Mooter G. Formulation and characterization of ternary solid dispersions made up of itraconazole and two excipients, TPGS 1000 and PVPVA 64, that were selected based on a supersaturation screening study.* Eur J Pharm Biopharm *2008;69:158–66.*

[n] *Vandercruys R, Peeters J, Verreck G, Brewster ME. Use of a screening method to determine excipients which optimize the extent and stability of supersaturated drug solutions and application of this system to solid formulation design.* Int J Pharm *2007;342:168–75.*

[o] *Sun Y, Tao J, Zhang GZ, Yu L. Solubilities of crystalline drugs in polymers: an improved analytical method and comparison of solubilities of indomethacin and nifedipine in PVP, PVP/VA, and PVAc.* J Pharm Sci *2010;99(9):4023–31.*

This additional testing may include BA studies depending on the development plan and timeline.

Dispersion selection is based on the characterization and property data collected for the dispersion samples. Early formulations may need certain properties that are essential in moving the molecule forward, such as improved solubility or BA, and these should be included in the selection criteria. A simple formulation with the necessary handling (such as temperature and low relative humidity) may be possible for small clinical supplies, but may not feasible for larger or longer clinical studies. Later in development, as the amount of clinical supplies increases and the formulation is modified, other processing, handling, and storage conditions would need to be considered. It is important to understand and include these development issues in the selection process.

Data obtained for the various dispersions can be reviewed in a number of ways. Two common methods are a decision tree[47] and a form matrix.[59] A decision tree narrows down the dispersion options based on a series of decision points. The dispersions that do not meet the criteria are removed from consideration and the rest are moved to the next decision point. Another option is a form matrix where the materials and properties are summarized in a table, acceptable properties are highlighted, and dispersions with the best properties are selected. For both approaches it is important to understand what is needed from an API property, formulation process, and development perspective. It should be noted that properties needed for early development may be different for a late-stage program and each selection should be assessed individually.

18.5 STABILITY CONSIDERATIONS

Amorphous materials are physically unstable and will have a tendency to crystallize to their more stable crystalline form (Fig. 18.2). One goal for amorphous solid dispersions is to maintain the amorphous API during drug product manufacturing and for the marketed drug's shelf life (physical stability). The second goal is to maintain a supersaturated solution of the API once the amorphous dispersion is dissolved in the gastrointestinal (GI) tract (solution stability); if the drug crystallizes upon contact with biological fluids or shortly after then the solubility advantage of the amorphous state is lost. This is illustrated in Fig. 18.5.[60]

Physical stability, or crystallization rate, in amorphous solid dispersions is dependent on a number of factors, such as the crystallization tendency of the drug alone,[61] miscibility,[17,33] storage temperature,[62] water content,[20] and the amount or type of polymer used.[20,63] Many of the polymers and excipients added

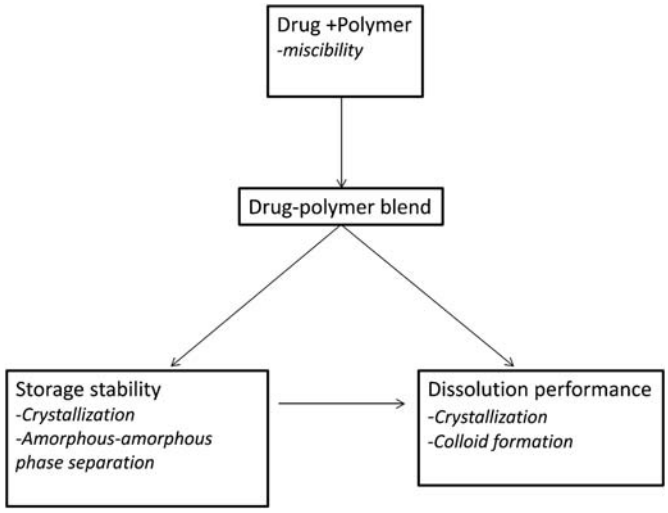


FIGURE 18.5 Schematic of factors important factors for the stability of amorphous solid dispersions. *Source: Reproduced with permission from Taylor LS. Chapter 5. Physical stability and crystallization inhibition. In:* Pharmaceutical amorphous solid dispersions. *A. Newman, editor. Hoboken, NJ: John Wiley and Sons; 2015.*

to amorphous solid dispersions can act as crystallization inhibitors and help maintain the amorphous nature of the drug in the solid phase. There are a number of possible stabilization mechanisms for polymers, including a decrease in molecular mobility related to T_g,[64,65] disruption of drug-drug interactions,[30] and formation of specific polymer-drug interactions related to miscibility.[32,66,67] It is likely that a combination of mechanisms contributes to the inhibition in most dispersion systems; however, these types of mechanisms can help guide polymer selection by comparing relevant API and polymer properties.

Assessing physical stability involves characterization methods (previously described) and presents a number of challenges. Physical stability testing is needed to determine which polymers best inhibit crystallization. This will involve different polymers, polymer grades, and concentrations, which can result in a large number of samples. Due to the small amount of API available in early development, these types of studies need to be carefully planned to obtain the most useful information. The low loading in amorphous dispersions also presents a challenge with many analytical techniques, and this aspect becomes more of an issue once the dispersion is formulated into a marketed product where the amount of active drug is further diluted. Determining crystallization rates as a function of temperature and water content can be difficult due to the lack of first principle models that accurately describe nucleation, growth, and overall crystallization kinetics in amorphous dispersions.[60] However, performing these studies may still provide a starting point for understanding the stability of a dispersion.

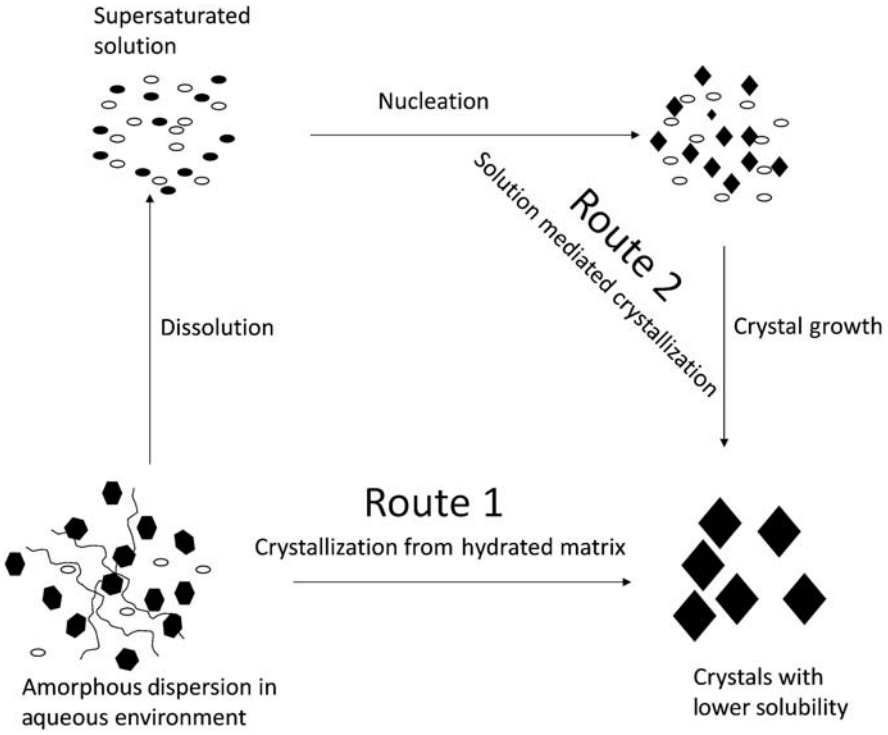


FIGURE 18.6 Schematic showing the possible routes of crystallization after adding an amorphous solid dispersion to an aqueous medium. Route 1 is matrix crystallization and Route 2 is solution crystallization. *Source: Reproduced with permission from Taylor LS. Chapter 5. Physical stability and crystallization inhibition. In:* Pharmaceutical amorphous solid dispersions. *A. Newman, editor. Hoboken, NJ: John Wiley and Sons; 2015.*

Crystallization from aqueous solutions, such as dissolution media or biological fluids in the GI tract, is another consideration. It is important that once the drug dissolves, the solution supersaturation is maintained to allow absorption of the drug. For dispersions with relatively high polymer content, the rate of release will likely be controlled by the polymer, not the API.[68] Fig. 18.6 shows two possible routes to crystallization when an amorphous solid dispersion is dissolved in an aqueous media: matrix crystallization (Route 1), where the dispersion may sorb water and the drug crystallizes in the matrix before dissolution, and solution crystallization (Route 2), where the drug is released into solution and then nucleates/crystallizes. For these types of studies, it is important to determine if crystalline or amorphous material is produced. If crystalline material is produced, the supersaturation of the solution will be lost, whereas if an amorphous material is produced, then supersaturation can be maintained.[69,70]

The ability of polymers or excipients to inhibit crystallization may influence precipitation (where turbidity is measured and it is not known whether crystalline or amorphous material is produced),[71] crystal nucleation,[72] or crystal growth.[73,74] In early development, studies can be performed to quickly assess the ability of various polymers to inhibit precipitation using turbidity[75] and can be performed in well plates.[76,77] A better understanding of the inhibition may be needed and more complex studies can be performed to obtain further information on the system as the drug progresses into development.

18.6 SOLUBILITY AND DISSOLUTION CONSIDERATIONS

Solubility and dissolution are important properties for amorphous solid dispersions because they have been linked to performance, such as BA. It has been shown that the flux across a membrane is enhanced for supersaturated solutions, but not for solubilized solutions;[78,79] therefore, using amorphous solid dispersions to create and maintain supersaturated solutions could lead to enhanced performance.

Solubilization depends on a number of drug properties, such as pKa, lipophilicity, and intermolecular forces in the solid state, as well as solvent properties, including pH, polarity, and solvent interactions. Solubility is a thermodynamic property and is defined by the concentration in solution in equilibrium with excess solid, therefore the value is based not only on the concentration in solution, but also by the solids present during the experiment. It is important to determine if equilibrium is established with an amorphous solid present during the measurement, with a crystalline material that has been produced during the solubility experiment, or colloidal drug aggregates.[80] The reference solubility is typically obtained with the most stable crystalline form. The solubility of other

TABLE 18.3 Method for Determining Solubility for Amorphous Solid Dispersions

Step	Comment
1. Ensure material has a uniform particle size to avoid multiple dissolution rates 75–150 μm is recommended	In early development this may difficult to obtain based on sample size: a range can be estimated using optical microscopy or SEM
2. Add a known excess of amorphous to medium (at least 10 × the crystalline solubility is recommended)	Amount of medium will depend on solubility and amount of sample available. Stirring (overhead or stir bar) can be limited by the volume of medium used
3. In order to improve powder dispersion, a paddle speed of at least 300 rpm is recommended	If wetting is not an issue, a lower speed, such as 100 rpm, is recommended
4. Draw sample at select time points and filter using an appropriate filter	Multiple time points are used to determine whether the profile levels off (equilibrium) or decreases (likely precipitation or crystallization)
5. Dilute samples immediately using organic solvent suitable for analysis method to prevent precipitation and/or crystallization	
6. Determine concentration in solution	
7. Pull an unfiltered sample at each time point and analyze for appearance of crystallinity	Wet sample analyzed by XRPD will provide information on the solid formed in solution; drying the sample may promote a form conversion and confuse the results
8. Determine the concentration versus time profile and note the first appearance of crystalline material	

Adapted from Murdande SB, Pikal MJ, Shanker RM, Bogner RH. Aqueous solubility of crystalline and amorphous drugs: challenges in measurement. Pharm Dev Technol *2011;16(3):187–200.*

metastable forms and supersaturation measurements reference that value.[81] In the case of an amorphous solid dispersion, the term *apparent solubility* is used to differentiate the amorphous value from the most stable crystalline form.[50,51,82] A suggested method for determining solubility is given in Table 18.3, which includes analysis of both the solution and the remaining solid.[83]

It is also important to decide which media will provide the most useful information. Solubility in water will provide a relatively straightforward comparison between the amorphous solid dispersion and a crystalline reference material This comparison can be used for initial assessment or for dispersion selection. However, it will not provide significant information on the solubility increase in biological media, which can be different based on pH and other factors. If absorption is expected in the intestine, then a corresponding media, such as simulated intestinal fluid, may be more appropriate. The transit time through the GI tract should also be considered when reviewing the time course of the experiment. If crystallization does not occur for 24 hours and the estimated residence time in the intestine is only a few hours, then the dispersion would be considered at low risk for crystallization.

Dissolution is different than solubility and applies to the kinetics rather than the thermodynamics of the process. Dissolution rate will increase with increasing surface area of the particles and drug solubility, and decrease with an increase in the thickness of the aqueous boundary and the concentration of drug already in solution. Dissolution will drive the initial solubilization of the dispersion and intestinal absorption will remove the solubilized drug, causing additional solubilization of the dispersion.

Typical dissolution methods based on United States Pharmacopeia (USP) criteria are used to determine similarity between lots for quality control purposes.[84] These methods are not usually applicable for evaluating in vivo performance. There are a number of parameters that need to be considered when developing a dissolution method for in vivo correlation, which are summarized in Table 18.4.[50] When developing dissolution methods for dispersions, the conditions chosen should mimic the GI tract, such as biorelevant media that matches the delivery route, pH, stirring rate, and volume. Most dissolution methods are run under sink conditions, however, nonsink conditions should be explored for poorly soluble drugs that may crystallize in the GI tract. Polymer properties also need to be considered and issues such as gelling, floating, or solids clinging to paddles, shafts, and glassware may be observed during method development. Crystallization propensity in the media over biologically relevant time frames should also be explored. For some dispersions, issues with wetting are overcome using surfactants.[26]

While Apparatus 1 (basket method) and Apparatus 2 (paddle method) are the most common USP methods used, it has been shown that Apparatus 3 (reciprocating cylinder) and Apparatus 4 (flow-through cell) offer advantages for low solubility compounds by allowing controlled pH and volume changes of the dissolution media throughout the analysis. A number of studies report the use of two-phase dissolution systems that attempt to mimic the GI tract (stomach and duodenum) and provide a better correlation with BA studies.[85–87] In one study, the crystallization of an amorphous form during dissolution was observed and

TABLE 18.4 Considerations for Amorphous Solid Dispersions[50]

Area	Considerations	Recommendations
Physical	Choice of polymer	Polymer characteristics (solubility, melting point, wettability, hygroscopicity, effect of pH, dissolution rate, etc.) and large-scale manufacturing process (spray drying, melt extrusion) need to be considered early in the development of the amorphous solid dispersion
	Polymer:API ratio	Optimize the amount of polymer required to provide the long-term physical stability desired in the solid state and crystallization inhibition in solution
	Miscibility	Determine if a one-phase miscible system has been produced
	Manufacturing processes	Consider that the amorphous solid dispersion produced at elevated temperatures may be miscible but may separate upon cooling to room temperature or under other conditions
	Hygroscopicity	Water uptake and the effect on T_g, physical stability, and crystallization need to be evaluated
Dissolution	Dissolution apparatus and media conditions	Dissolution method development should focus on conditions that mimic the GI tract (biorelevant media that match the delivery route, pH, stirring rate, volume); these conditions may not be applicable for routine dissolution testing to discriminate differences between lots
	Sink versus nonsink conditions	Both sink and nonsink conditions should be investigated and compared with data from the in vivo studies; for poorly soluble drugs nonsink conditions and the possibility of crystallization in the GI tract need to be evaluated
	Polymer controlled dissolution/wettability	Incorporate polymer properties into the dissolution method development (gel formation of the polymer, floating of particle/dispersion particles, solid clinging to paddles/shafts/glassware)
	Supersaturation in dissolution testing over hours	Test supersaturation over biologically relevant time frames in biorelevant media to determine if the API will stay in solution or crystallize out
Biological	Fasting versus Fed-pH effects	Food effects need to be evaluated; specifics on fasting times and post-dose feedings need to be detailed in reported studies
	Sink versus nonsink conditions	Dose-range dependent sink conditions along the GI tract (pH, volumes, residence times) need to be considered and investigated; these will be dependent on the solubility of the API and the amorphous solid dispersion
	Local effect of polymers on pH	pH effects of polymers, especially polymers used for enteric coatings, need to be evaluated throughout the entire GI tract
	Species differences	Investigate the use of alternative animal models, such as pigs and minipigs, that are a better model for non-primate studies
	Transporters and metabolism	Determine if APIs are substrates for human transporters and/or metabolizing isoforms and the affect this will have on absorption
	Gastrointestinal physiology	For highly lipophilic molecules, lymphatic absorption and not only blood absorption need to be determined in early studies; use extended GITA model to predict regions of optimal absorption; use feedback from in silico models to help refine formulations

dissolution results compared favorably with dog BA studies.[85] Fed versus fasted conditions can also be simulated with the media chosen for the dissolution method and this factor should be considered, and be similar to what will be used in the animal studies, when developing the method.

18.7 METHODS OF MANUFACTURING AMORPHOUS SOLID DISPERSIONS

A variety of methods are used to produce amorphous solid dispersions at small scales, but the most common methods to produce large amounts of material (hundreds of grams or kilogram quantities) are hot melt extrusion[88] and spray drying.[48] Another method that uses common formulation processing equipment to produce large-scale quantities of dispersions, fluid bed granulation, has also been used for certain drugs. All three methods have been used to produce marketed products, as summarized in Table 18.1.

Melt extrusion forms a dispersion by melting the API/polymer mixture followed by rapid cooling to solidify the components and prevent API crystallization (Fig. 18.7). A variety of polymers can be used for this process[49] and the most common polymers include

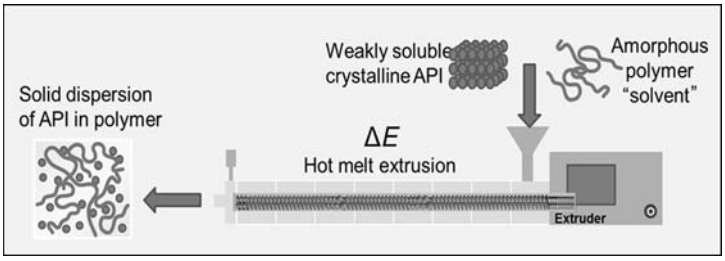


FIGURE 18.7 Hot melt extrusion process schematic. *Source: Reproduced with permission from Crowely K, Gryczke A. Chapter 10. Hot melt extrusion of amorphous solid dispersions. In:* Pharmaceutical amorphous solid dispersions. *A. Newman, editor. Hoboken, NJ: John Wiley and Sons; 2015.*

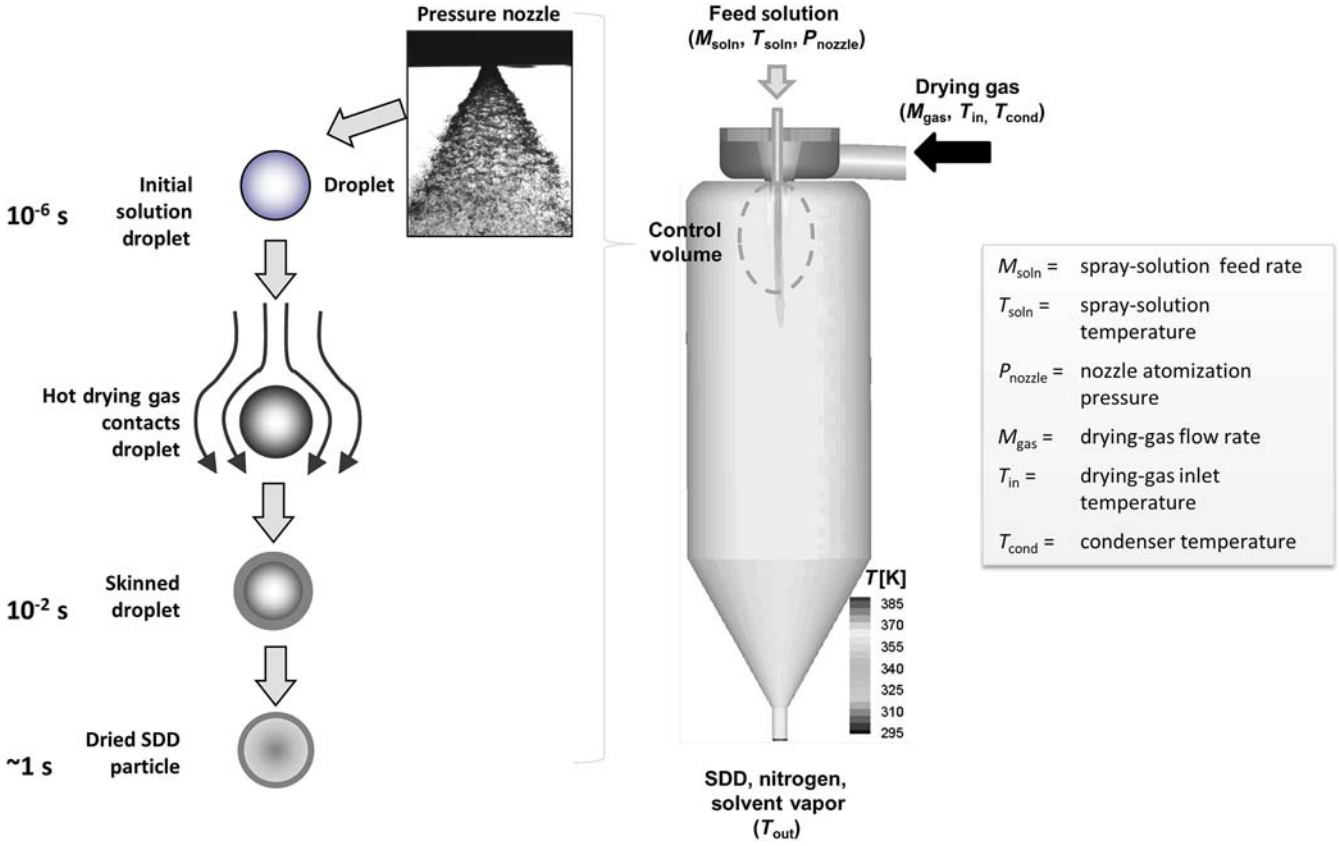


FIGURE 18.8 Spray-drying process schematic. *Source: Reproduced with permission from Dobry DE, Settell DM, Bauman JM. Chapter 9. Spray drying and scale-up. In:* Pharmaceutical amorphous solid dispersions. *A. Newman, editor. Hoboken, NJ: John Wiley and Sons; 2015.*

copovidone, povidone, HPMCAS, polyethylene oxide, polyethylene glycol (PEG), HPMC-E5, and methacrylates. The melting and T_g values of both drug and polymer are used to determine processing temperatures, usually 20−50K greater than the polymer T_g. Plasticizers, such as PEGs, poloxamers, and triethyl citrate, may be added to lower thermal transitions of the individual components or the extruded dispersion.[89] Additives to alter dissolution, such as surfactants, may be included to extend the duration of supersaturation or delay drug release. Extruders usually run in a continuous mode. Key process parameters include extruder screw design, feed rate, and temperature. It is critical that the API and excipients do not degrade at the processing temperature. The extrudate undergoes downstream processing to tailor the physical form for the final dosage form. Downstream processing can include pelletization or milling.

Spray drying is a solvent-based method commonly used to produce amorphous solid dispersions where droplets are produced and the solvent is removed within milliseconds (Fig. 18.8).[90] It is important to obtain solubility values of both the API and polymer in a solvent suitable for spray drying, such as dichloromethane, acetone, methanol, tetrahydrofuran, ethanol, ethyl acetate, water, and mixtures of these solvents. Both the API and polymer are dissolved in the solvent system. The spray-drying process consists of four stages: atomization of the liquid, mixing of the liquid with the drying gas, evaporation of the liquid, and

separation of the dried particles from the gas. By controlling various parameters, such as the atomization and drying parameters, dispersion particles can be engineered with specific properties, such as particle size and density. Secondary drying, using tray drying or vacuum drying, is added in many systems to remove residual solvent from the spray-dried dispersions. Small-scale spray dryers are available for milligram quantities of material, laboratory-scale equipment can produce gram-scale quantities, and production scale dryers produce metric ton batches. Spray drying has demonstrated scalability over these various batch sizes.[48,90]

Amorphous dispersions can also be produced using conventional manufacturing equipment, such as a fluid bed dryer/granulator or high-shear mixer, by spraying an API/polymer solution onto an inert substrate. The marketed product Sporonox is produced by a spray-layering process where a round core is initially sprayed with a coating film, followed by an amorphous dispersion layer containing HPMC and itraconazole, and then a seal-coating layer of PEG to prevent sticking of the beads.[91] These beads are then filled into capsules for the marketed product. Other methods have been used to produce dispersions at a scale for marketed products, such as melt onto a carrier (Afeditab) and antisolvent addition (Zelboraf). Detailed discussion of basic principles, process development, and scale-up of these manufacturing technologies can be found in relevant chapters in this book.

18.8 DOSAGE FORM DEVELOPMENT CONSIDERATIONS

Amorphous solid dispersions have been used in a variety of marketed products, as shown in Table 18.1. Once the amorphous solid dispersion has been chosen, it is important to understand how the large-scale production parameters may be related to the physical properties of the dispersion influencing dissolution, such as particle size and bulk density. Often, the manufacturing process will need to be examined and modified to produce large-scale batches with the desired properties.

Dispersion properties also need to be considered for the formulation process for a solid oral dosage form. Melt extruded dispersions have a tendency to be more dense, which may favor capsule formulations but may exhibit poor compression properties for tablets; therefore, a method to decrease the density may be needed. Spray-dried dispersions may consist of fine particles, exhibit poor flow and compression properties, and be poor candidates for capsule formulations. Methods to improve flow and compressibility of the spray-dried materials may be needed before tableting.

The physical stability of the dispersion will influence handling and storage conditions. For early formulations, the dispersion may be dispersed as a slurry in a liquid formulation vehicle. In correlating to BA, it is critical to determine how much API can be dissolved in the vehicle, as well as how long the solid remains amorphous in the liquid vehicle to determine a time frame for dosing after preparation. For dispersions that will be filled into capsules, the relative humidity of the room during encapsulation may need to be monitored and kept below a certain threshold to prevent crystallization. The amount of water in the capsules and possible migration to the solid amorphous dispersion placed in the capsule may need to be investigated.

The choice of excipients is a critical parameter to evaluate when using dispersions, especially in later development when the formulations become more complex. If the dispersion exhibits poor physical stability, especially around water, then anhydrous excipients should be considered. Some dispersions will show less issues with water. For Intelence, the package insert states that the tablet can be dispersed in water and taken immediately as a slurry.[92]

Formulation processes need to be considered because compression, heat, and exposure to moisture are commonly used to produce the final dosage form. Determining methods to measure physical stability (crystallization) and chemical stability (degradation) will be a key step in producing viable formulations. Analysis of the solid at different stages of the process will help narrow problem areas that would need remediation.

Stability testing of the drug product is needed to determine packaging, storage, and shelf life. The Food and Drug Administration has issued guidelines for stability conditions and time periods needed for long-term, intermediate, and extended stability studies on drug products.[93] Data must be examined to determine the conditions that affect the stability of the drug, and the process and packaging will need to be modified to ensure that the API remains amorphous.

18.9 CASE STUDIES

As discussed in this chapter, amorphous solid dispersions in early- and late-stage development will have different requirements, and the types of studies needed at each stage will be tailored as the product's launch approaches. In this section, case studies involving amorphous solid dispersions in early and late development, as well as in life cycle management, will be presented.

18.9.1 Early Development: Vemurafenib

Vemurafanib (Fig. 18.9) is a compound that is practically insoluble at pH 3 and 7 (<0.1 μg/mL), and an amorphous solid dispersion was used to improve its solubility. In the case of vemurafanib, a very high melting point and very low solubility in low-boiling solvents made conventional spray drying and melt extrusion methods unsuitable. The dispersions were prepared by a coprecipitation process shown in Fig. 18.10, which is an alternative preparation method that can be used when spray drying or melt extrusion are not feasible.[94] Dispersion samples prepared by the coprecipitation method were called microprecipitated bulk powders (MBP). Enteric polymers HPMCAS, hydroxypropyl methylcellulose phthalate (HPMCP), and Eudragit L 100-55 in a 40:60 or 30:70 w/w% drug: polymer ratio were used in the study. A solution of drug and stabilizing polymer in dimethylacetamide (DMA) is added to a dilute hydrochloric acid (HCl) solution maintained at 2–5°C; precipitation occurs

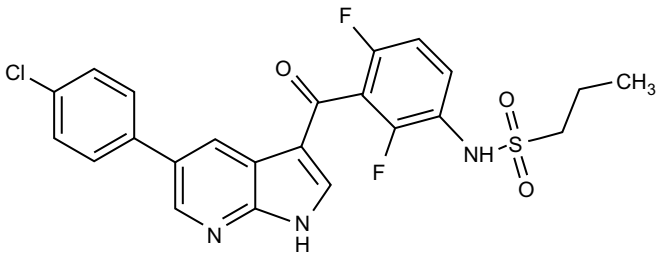


FIGURE 18.9 Structure of vemurafanib.

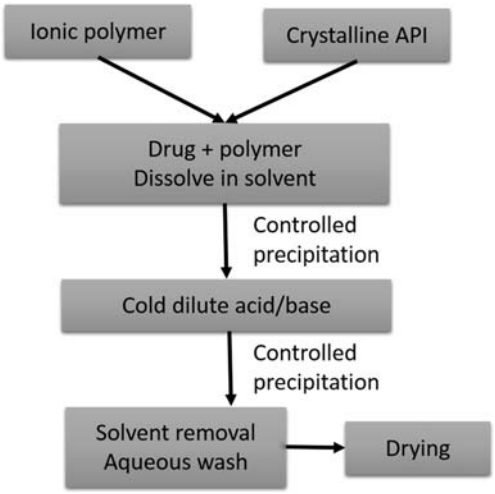


FIGURE 18.10 Schematic of the coprecipitation method used to produce MBP samples. *Source: Adapted from Shah N, Iyer RM, Mair H-J, Choi DS, Tian H, Diodone R, et al. Improved human bioavailability of vemurafenib, a practically insoluble drug, using an amorphous polymer-stabilized solid dispersion prepared by a solvent-controlled coprecipitate process.* J Pharm Sci *2013;102:967–81.*

based on the insolubility of the drug and polymer in acidic aqueous media, resulting in the API being dispersed in the polymer matrix.

Characterization of the samples showed that all three dispersions were amorphous. Upon storage at 40°C/75% RH in open containers for 4 weeks, the HPMCP dispersion exhibited signs of crystallization after 2 weeks while the other two polymers exhibited no crystallization after 4 weeks. Based on this, the HPMCP dispersion was not pursued further. One T_g value was measured for the remaining dispersions (102°C for the HPMCAS system and 111°C for the Eudragit system), suggesting a miscible system and good physical stability at room temperature (RT). Dissolution studies showed that HPMCAS was able to maintain a higher level of supersaturation (90%) than Eudragit (55%) over a 5-hour time frame. These values represent supersaturated concentrations levels of approximately 40 μg/mL (HPMCAS) and 22 μg/mL (Eudragit), which is almost a 100-fold increase for the HPMCAS dispersion over the crystalline drug. A pH solubility study ranging from pH 2 to 8 resulted in a two–threefold greater solubility in the enteric pH range of 6.8–8.0 for the HPMCAS material over the Eudragit L 100-55. Based on the higher supersaturation potential, the HPMCAS dispersion was chosen for further development.

Additional studies on the HPMCAS dispersion were performed. Scanning electron micrography (SEM) micrographs showed that the MBP particles have a relatively smooth and round surface, whereas the inner structure is a spongy network with pores in the range of 50–200 nm and larger bubbles in the range of 3–10 μm. Crystalline particles related to crystalline API were not detected in SEM images, indicating a homogenous dispersion of amorphous API and polymer, which agrees with the XRPD and DSC data. Dissolution studies using a flow-through cell showed that 100% of drug was released from the MBP after 120 minutes, compared with about 10% release from amorphous drug and a physical mixture of crystalline vemurafenib and HPMCAS. The particle size distribution of the MBP samples exhibited a d_{50} value of 60 μm and a d_{90} value of 220 μm. The dissolution rate of MBP samples of different particle sizes showed a decrease in dissolution with larger particles (125 μm) compared with smaller particles (48 μm), indicating that dissolution could be further enhanced by optimizing the particle size distribution.

HPMCAS MBP samples stressed at 40°C/100% RH showed a partial conversion to crystalline API (Fig. 18.11a) based on an XRPD quantitative method able to detect crystalline levels as low as 1.6%. Dissolution profiles of stressed and unstressed MBP samples (Fig. 18.11b) resulted in lower dissolution

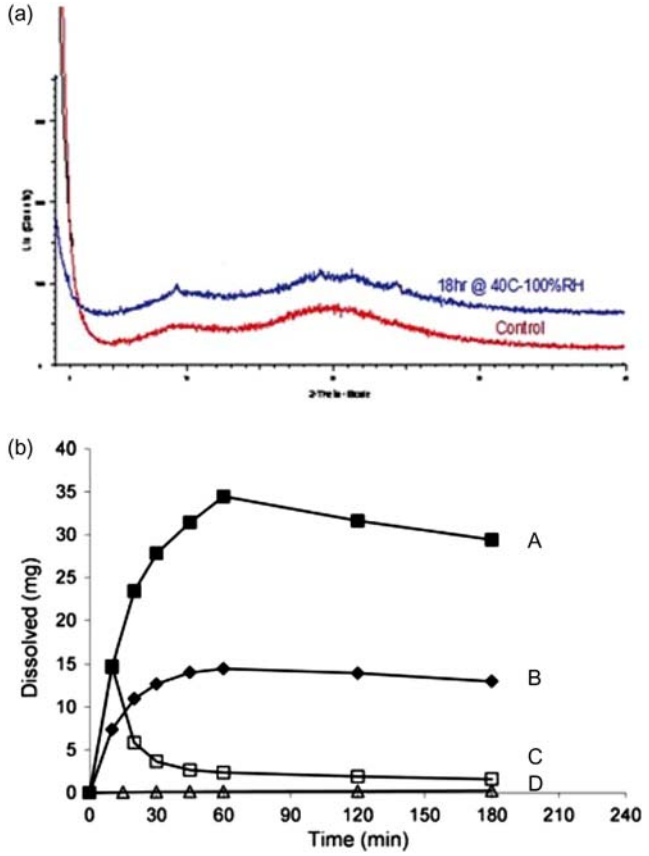


FIGURE 18.11 (a) XRPD profiles of control and stressed capsules of vemurafenib:HPMCAS MBP; (b) dissolution profiles of vemurafenib:HPMCAS MBP and crystalline vemurafenib: (A) unstressed vemurafenib:HPMCAS MBP, (B) stressed vemurafenib:HPMCAS MBP, (C) metastable crystalline vemurafenib, (D) stable crystalline vemurafenib. *Source: Reproduced with permission from Shah N, Iyer RM, Mair H-J, Choi DS, Tian H, Diodone R, et al. Improved human bioavailability of vemurafenib, a practically insoluble drug, using an amorphous polymer-stabilized solid dispersion prepared by a solvent-controlled coprecipitate process.* J Pharm Sci *2013;102:967–81.*

rates for the stressed samples due to the increased levels of crystalline material. A correlation plot of percent crystalline material and dissolution rate showed a good relationship between the two parameters (coefficient of correlation = 0.9753). Upon storage, the water content of MBP samples with different initial moisture contents (<1% vs approximately 2%) was found to be about the same, reaching a final water content of approximately 4% within 2 months of open storage at 30°C/75% RH and 40°C/75% RH. These data were used to determine handling and storage conditions of the MBP to maintain the amorphous nature of the solids.

The HPMCAS MBP was formulated into 40-mg strength capsules of vemurafenib, employing dry blending (MBP-1) and wet-mixing (MBP-2) processes. The effect of temperature and humidity on moisture content and percent crystallinity was evaluated after

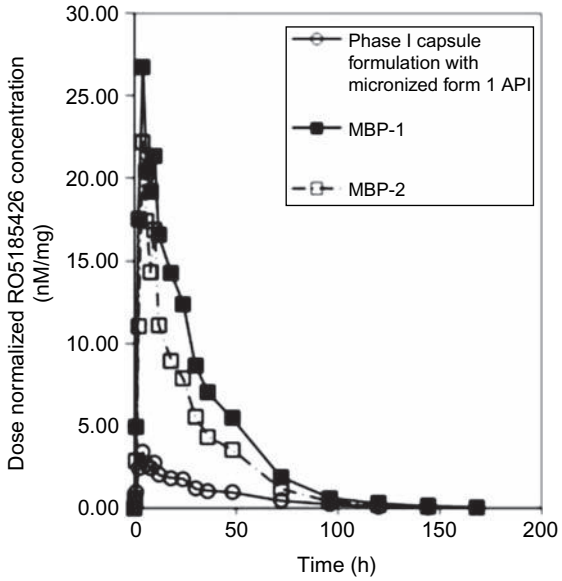


FIGURE 18.12 Comparison of dose-normalized exposure data among three capsule formulations: Phase I crystalline vemurafenib and two MBP amorphous vemurafenib formulations. *Source: Used with permission from Shah N, Iyer RM, Mair H-J, Choi DS, Tian H, Diodone R, et al. Improved human bioavailability of vemurafenib, a practically insoluble drug, using an amorphous polymer-stabilized solid dispersion prepared by a solvent-controlled coprecipitate process.* J Pharm Sci *2013;102:967–81.*

15 months using open and closed storage conditions (25°C/60% RH, 30°C/75% RH, and 40°C/75% RH). The open conditions resulted in water contents ranging from 3.5% to 4.5% and crystalline contents ranging from 2% to 24%. Closed conditions showed significantly less water (1.3–1.6%), and only 2% crystalline material under the most severe stress conditions. Based on these data, it was determined that protective packaging was essential to control moisture and maintain the amorphous nature of the HPMCAS dispersion.

A relative BA study in healthy volunteers was conducted using MBP-1, MBP-2, and a 100-mg Phase I crystalline drug capsule formulation containing 10% poloxamer F127. The study was conducted for 84 days with 18 healthy, male subjects ranging in age from 18 to 65 years old. Both MBP formulations had nearly identical exposures following a single 160-mg dose, with area under curve (AUC) values of 86 (MBP-1) and 80 μ Mh (MBP-2), compared with a value of 34 μ Mh for a single 300-mg dose of the Phase I crystalline formulation (Fig. 18.12). The relative BA for the MBP-1 and MBP-2 formulations was 4.4–4.7 times compared with the crystalline formulation. The maximum concentration (C_{max}) ratios were also found to be similar for both MBP formulations (3.6 and 3.1) when compared with the crystalline formulation. Based on these data, it was evident that the MBP formulations were superior to the Phase I crystalline formulations and

supported further development of the amorphous dispersion to improve the BA of vemurafenib.

This case study uses a bulk precipitation method that can be used when conventional spray drying and melt extrusion methods are not feasible. This preparation can be used at small or large scale, enabling relatively easy production of larger batches throughout the development process. A number of studies were needed on both the MBP and formulated product to understand the physical stability and effect of crystallization on dissolution, as well as to determine handling and packaging conditions. The result was an amorphous dispersion capsule formulation with improved BA at lower doses, which was a significant improvement over the Phase I crystalline formulation.

18.9.2 Late Development: Telaprevir

Telaprevir (Fig. 18.13) is small molecule that targets the hepatitis C virus. The molecule is an active site inhibitor and needs to be apolar in order to bind effectively, which then results in a very low aqueous solubility. Early development studies determined that an amorphous dispersion was needed to obtain the necessary properties for oral administration. An amorphous dispersion with HPMCAS was chosen for further development.[95]

After early development studies were completed, the amorphous solid dispersion needed to be formulated into clinical and commercial drug products. A spray-drying process was chosen for large-scale manufacturing of the dispersion based on the acceptable solubility of telaprevir in organic solvents and the ease of scaling up the process.[96] The overall formulation process would include secondary drying to remove excess residual solvent, blending with excipients, compressing the blend into tablets, and film coating the tablets. A quality by design (QbD) approach was used for late-stage development. The quality target product profile for the commercial drug product included (1) an immediate release tablet; (2) stable during 24 months of shelf storage; (3) orally administered; (4) bioavailable; (5) safe; (6) efficacious; and (7) easily distinguishable from other medications. From this profile, the critical quality attributes (CQAs) for the dispersion and the coated tablets were defined, and the possible impact on quality, safety, or efficacy were assessed (Table 18.5).

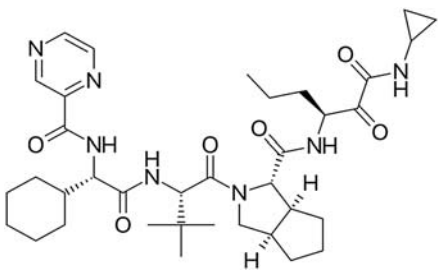


FIGURE 18.13 Structure of telaprevir.

The initial risk assessment suggested that the dispersion is a potential critical material due to its impact on dissolution. Each unit operation was studied to determine possible critical and key processing parameters using a design of experiment (DoE) approach. Table 18.6 lists dispersion manufacturing and tablet preparation steps that were identified to affect the dissolution. Based on this list, a set of experiments were identified to define the process design space for tablet CQA of dissolution. Dispersion attributes that needed to be controlled as part of the spray drying and secondary drying process were identified as particle size, bulk density, and wet solid form prior to secondary drying (must be amorphous). Thirteen factorial experimental designs were conducted, covering 148 runs, to evaluate factors that influenced the spray-drying process. An additional 14 runs were added from data obtained from the initial runs. The spray-drying process space was found to rely on (1) wet particle size, wet bulk density, and spray drying process parameters and (2) the change in particle size and bulk density after secondary drying. The acceptable ranges for particle size and bulk density were found to be $50-110\,\mu m$ and $0.33-0.52\,g/mL$, respectively, and spray drying conditions were found to consistently provide material with these specifications.

The spray-drying parameters were then used as inputs for a DoE on tablet compression (Fig. 18.14). A model was developed to estimate tablet dissolution based on the dispersion particle size, bulk density, and average tablet hardness. To produce the dissolution models, 241 tablet sublots were employed, and the mean cumulative release of drug from the tablet after 15 minutes (normalized to the 90 minutes point) was used to determine the factors affecting the dissolution. The early timepoint was found to be more sensitive to processing and dispersion particle size and bulk density. A value of $>80\%$ dissolution at 15 minutes was used to define the design space for compression and to ensure that the dissolution specification (Q at 45 minutes = 80%) was met. A mean average tablet hardness of 13 kP was assumed for the next stage, which was to reduce the model to a relationship between particle size and bulk density, representing the limiting design space to achieve acceptable dissolution.

The next phase was a QbD study on the physical and chemical stability of the drug product to determine shelf life. A temperature and relative humidity study was included to not only assess physical stability, but also to develop a chemical stability model that correlated the amount of epimer (the primary degradant of the API) with factors such as surface pH of excipient, moisture content of the tablet, storage temperature, and

TABLE 18.5 Summary of Critical Quality Attributes for Telaprevir[96]

Quality attributes	Spray-dried dispersion	Drug product	Potential impact on quality, safety, and/or efficacy
Identification	X	X	Ensures correct active ingredient
Appearance	X	X	Visual indicator of product quality
Assay	X	X	Ensures potency of subsequent tablet
Degradation products	X	X	Ensures that levels of impurities do not affect assay or safety (by exceeding the levels deemed safe in toxicology or clinical studies)
Residual solvents	X		Affects safety, potency, and physical stability of the product
Water content	X	X	Ensures that water content is not so high as to lower chemical stability or impact physical stability during the product's shelf life
Physical form	X[a]	X[a]	Ensures that telaprevir remains amorphous and, therefore, bioavailable
Particle size	X		Affects flowability and compactability of the powder during tableting which could lead to low potency or broken tablets (efficacy). Incorrect particle size affects downstream manufacturability through inability to meet content uniformity, tablet weight, and/or hardness. Particle size may impact dissolution
Powder bulk density	X		Affects flowability and compactability of the powder during tableting which could lead to low potency or broken tablets (efficacy). Incorrect bulk density affects downstream manufacturability through inability to meet content uniformity, tablet weight, hardness and/or appearance. These variations could have an impact on tablet dissolution
Tablet hardness		X	Ensures that the integrity of tablets is not compromised
Dose uniformity		X	Ensures dose is uniform across the batch
Dissolution		X	Impacts the rate of drug release
Microbial limits		X	Ensures safety of the drug product

[a]*Physical form should be amorphous for dispersion and tablet.*

TABLE 18.6 List of Parameters Identified to Affect Dissolution of Vemurafenib Drug Product[96]

	SDD process steps		Tablet process steps		
	Spray drying	**Secondary drying**	**Blending**	**Compression**	**Film coating**
Parameters	Equipment Solids load Nozzle diameter Feed pressure Outlet T Condenser T Drying gas flow rate	Drying T N2 flow rate % Fill volume Wet SDD property Agitation rate	Lube blend revolution	SDD PS SDD BD Main compression force	Coating weight gain
Output to evaluate	PS BD Physical form	PS BD Physical form	% RSD	Hardness Dissolution Friability Thickness Appearance	Dissolution Water Content

storage time. Tablets were equilibrated at conditions ranging from 10% to 60% RH, and using tablet moisture contents ranging from 0.4% to 3.4%; the tablets were then sealed into foil pouches to maintain constant water content. The pouches were exposed to 25°C, 30°C, and 40°C for up to 12 months. A regression model was developed to predict the growth of epimer and was used to determine appropriate acceptance criteria for water content in the tablets and impurity content in an excipient. Good agreement was found with traditional stability data and the model was able to predict the amount of epimer formed.

Based on the epimer data, the maximum allowable tablet water content and maximum epimer value

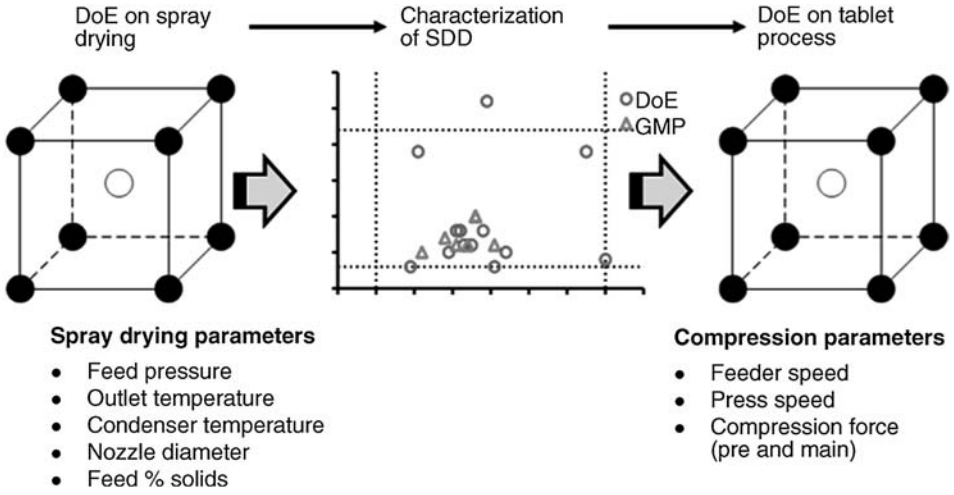


FIGURE 18.14 Interaction between outputs of spray drying and tablet compression. *Source: Reproduced with permission from Sanghvi T, Katstra J, Quinn BP, Thomas H, Hurter P. Chapter 11. Formulation development of amorphous dispersions. In Pharmaceutical amorphous solid dispersions. A. Newman, editor. Hoboken, NJ: John Wiley and Sons; 2015.*

(4.5%) during storage for 24 months at 25°C/60% RH was determined, and two packaging regimes were tested. The first, impermeable packaging, needed a maximum tablet water content of 2.4% to keep the epimer content less than 2% during a 24-month shelf life at 25°C/60% RH, and a tablet water content of 2.1% for storage at 30°C/60% RH for 12 months. Permeable packaging required a moisture vapor transmission rate (MVTR) less than 0.149 mg/day/tablet for a tablet containing 1.5% water and stored at 25°C/60% RH for 24 months and below 0.137 mg/day/tablet for storage at 30°C/60% RH for 12 months. These data showed that a maximum tablet water content of 1.5% at release and 2.5% on stability, along with the packaging having a maximum MVTR of 0.137 mg/day/tablet would give the necessary shelf life.

This case study illustrates the work needed during late development to produce a late-stage product that is suitable for market. A QbD approach was used to link spray-drying conditions, dispersion attributes, and tablet compression properties to dissolution performance. Understanding the interrelationships between all of these parameters was needed to produce a stable amorphous dispersion with acceptable properties. Once the optimal formulation was produced, additional work was performed to determine packaging that would maintain chemical stability over a 2-year shelf life.

18.9.3 Life Cycle Management

18.9.3.1 Kaletra

Kaletra is combination product containing ritonavir and liponavir, both biopharmaceutics classification system (BCS) 4 compounds, used to treat AIDS.

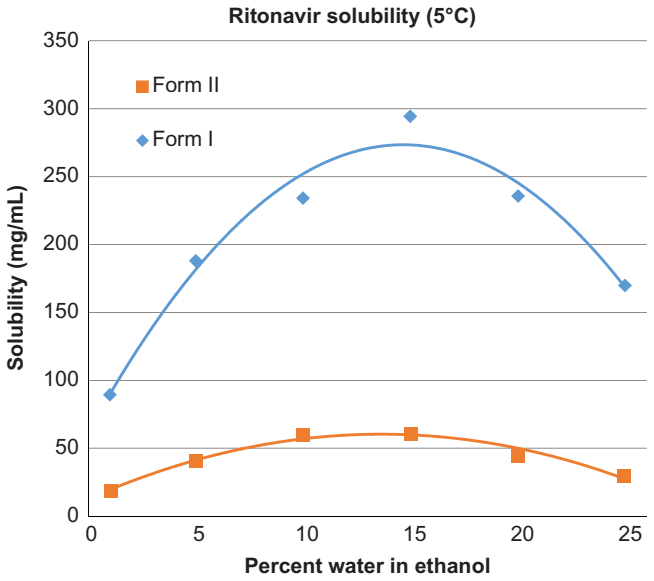


FIGURE 18.15 Ethanol:water solubility of ritanovir Forms I and II. *Source: Adapted from data in Bauer J, Spanton S, Henry R, Quick J, Dziki W, Porter W, et al. Ritonavir: an extraordinary example of conformational polymorphism.* Pharm Res *2001;18:859–66.*

Development started with ritonavir, which was discovered in 1992. It was not bioavailable from the solid state and not water soluble. The marketed product (Norvir) was available in 1996, with ritanovir dissolved in ethanol:water-based solutions to produce an oral liquid or a semisolid capsule formulation. Only one crystalline form was identified during development and in 1998, after 240 lots of capsules, a second form was identified in the capsules which was significantly less soluble in the ethanol:water system (Fig. 18.15).[97,98] The product was then reformulated with the new form.

Alternative methods to improve the solubility of ritonavir were explored, including amorphous solid dispersions.[99,100] Early studies investigated dispersions with PEG, and a significant increase in plasma concentration was observed for samples containing 10:90 ritonavir:PEG compared with crystalline material in dog studies. The 10% dispersion exhibited a 22-fold increase in AUC and a 13.7-fold increase in C_{max}. Lopinavir was a second HIV protease inhibitor under development and was found to be 10-fold more potent than ritonavir.[101] It was also found to be poorly soluble, and issues with low BA and fast elimination were observed when lopinavir was orally administered. It was found that coadministration with ritanovir dramatically improved the poor pharmacokinetic properties by preventing metabolism of liponavir. The combination of lopinavir and ritonavir produced a 77-fold increase compared with dosing with liponavir alone. Conventional solid tablet, capsule, and suspension formulations did not produce adequate plasma levels compared with soft-gel capsules,[102] and Kaletra soft-gel capsules were launched in 2000.

An amorphous solid dispersion of lopinavir and ritonvavir was investigated using melt extrusion technology, and a dispersion containing both drugs and copovidone was developed.[103–105] A number of processing parameters were investigated, such as mixing zone length (first, second, and overall), mixing zone angle, temperature, feed rate, and screw speed. Analytical results included amorphous nature, degradation products, and water content. A 200/50-mg lopinavir/ritonavir tablet formulation was developed that had a number of advantages over the soft gel formulation that was on the market. The soft gel was approved at a dose of 400/100-mg (three capsules) taken twice daily with food. Studies with the solid amorphous dispersion (200/50-mg lopinavir/ritonavir tablet formulation) resulted in more consistent lopinavir and ritonavir exposures, diminished food effect, and no refrigeration.[102] The plot for lopinavir and ritonavir BA under different fed/fasted conditions clearly showed no difference (Fig. 18.16).[106] Kaletra tablets were introduced in 2005.

This case study shows how changing the solid-state form from a crystalline material to an amorphous solid dispersion allowed development of a solid dosage form, which exhibited significant advantages over the initial soft-gel capsules. Melt extrusion was used to produce a dispersion that could be easily be formulated into a conventional tablet. A better product for patients was produced by decreasing the number of tablets, reducing food effects so it would not have to be taken with a meal, and storage at room temperature rather than refrigerated conditions.

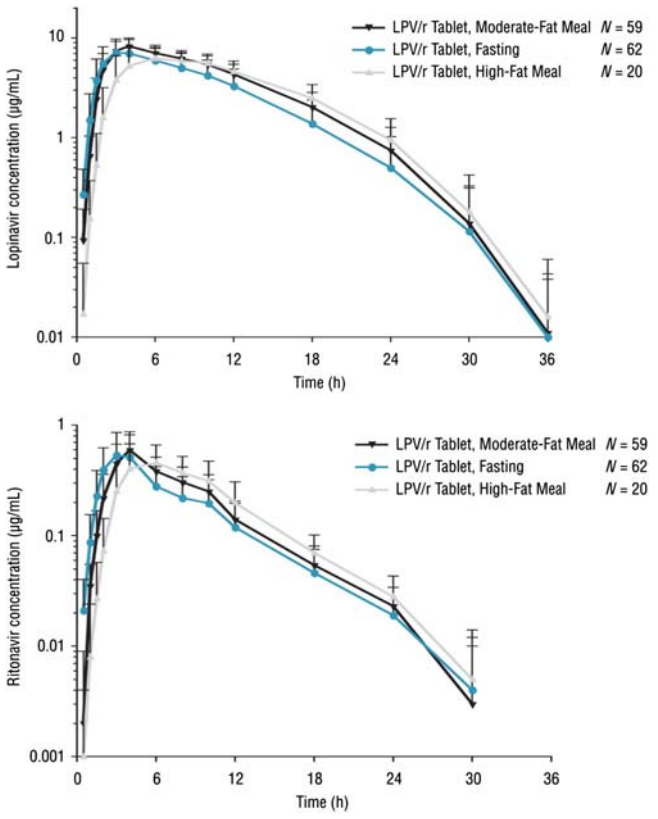


FIGURE 18.16 Plasma concentration of lopinavir (top) and ritonavir (bottom) after dosing Kaletra under different fed/fast conditions.[106]

18.10 CONCLUSIONS

Amorphous solid dispersions are one way to improve the properties of an API and they are typically used to help improve the solubility of poorly soluble compounds. They also have disadvantages, such as poor physical stability, which can result in crystallization, and poor chemical stability, which can lead to degradation. The screening and selection process needs to include knowledge of required dispersion properties as well as downstream processing to produce the formulated product. Since they are metastable forms, development of amorphous dispersions involves a greater understanding of the handling, storage, and processing conditions needed to prevent crystallization and to maintain a miscible dispersion. A number of marketed products containing amorphous solid dispersions show that this technology has been successful in developing poorly soluble compounds.

References

1. Byrn SR, Pfeiffer RR, Stowell JG. *Solid-state chemistry of drugs*. 2nd ed. West Lafayette, IN: SSCI, Inc; 1999.
2. Bates S, Zografi G, Engers D, Morris K, Crowley K, Newman A. Analysis of amorphous and nanocrystalline solids from their X-ray diffraction patterns. *Pharm Res* 2006;**23**(10):2333–49.

3. Shah N, Sandhu H, Choi DS, Chokshi H, Malick AW, editors. *Amorphous solid dispersions theory and practice*. New York: Springer; 2014.
4. Newman A, editor. *Pharmaceutical amorphous solid dispersions*. Hoboken, NJ: John Wiley and Sons; 2015.
5. Harmon P, Li L, Marsac P, McKelvey C, Variankaval N, Xu W. Amorphous solid dispersions: analytical challenges and opportunities. *AAPS Newsmagazine* 2009;Sept:14–20.
6. Tao J, Sun Y, Zhang GGZ, Yu L. Solubility of small-molecule crystals in polymers: D-mannitol in PVP, indomethacin in PVP/VA, and nifedipine in PVP/VA. *Pharm Res* 2009;**26**(4):855–64.
7. Chokshi RJ, Zia H, Sandhu HK, Shah NH, Malick WA. Improving the dissolution rate of poorly water soluble drug by solid dispersion and solid solution- pros and cons. *Drug Deliv* 2007;**14**:33–45.
8. Serajuddin ATM. Solid dispersion of poorly water-soluble drugs: early promises, subsequent problems, and recent breakthroughs. *J Pharm Sci* 1999;**88**:1058–66.
9. Chou WL, Riegelman S. Pharmaceutical application of solid dispersion systems. *J Pharm Sci* 1971;**9**:1281–302.
10. Vasanthavada M, Tong W-Q, Joshi Y, Kislalioglu MS. Phase behavior of amorphous molecular dispersions I: determination of the degree and mechanism of solid solubility. *Pharm Res* 2004;**21**(9):1598–606.
11. Lu X, Wenslow R, Newman A. Chapter 14. Monographs on polymers and surfactants. In: Newman A, editor. *Pharmaceutical amorphous solid dispersions*. Hoboken, NJ: John Wiley and Sons; 2015.
12. Nagarsenkar MS, Rupalin M, Ramprakash G. Solid dispersion of hydroxypropyl β-cyclodextrin and ketorolac: enhancement of in-vitro dissolution rates, improvement in anti-inflammatory activity and reduction in ulcerogenicity in rats. *J Pharm Pharmacol* 2000;**52**:949–56.
13. Law SL, Lo WY, Lin FM, Chaing CH. Dissolution and absorption of nifedipine in polyethylene glycol solid dispersions containing phosphatidylcholine. *Int J Pharm* 1992;**84**:161–6.
14. Newman A. Chapter 3. Screening and selection. In: Newman A, editor. *Pharmaceutical amorphous solid dispersions*. Hoboken, NJ: John Wiley and Sons; 2015.
15. Lim H-T, Balakrishnan P, Oh DH, Joe KH, Kim YR, Hwang DH, et al. Development of novel sibutramine baase-loaded solid dispersion with gelatin and HPMC: physicochemical characterization and pharmacokinetics in beagle dogs. *Int J Pharm* 2010;**397**:225–30.
16. (a) Lobmann K, Grohganz H, Laitinen R, Strachan C, Rades T. Amino acids as co-amorphous stabilizers for poorly water soluble drugs—Part 1: preparation, stability, and dissolution enhancement. *Eur J Pharm* 2013;**85**(3):873–81.(b) Lobman K, Laitinen R, Strachan C, Rades T, Grhganz H. Amino acids as co-amorphous stabilizers for poorly water-soluble drugs—Part 2: molecular interactions. *Euro J Pharm Biopharm* 2013;**85**(3):882–8. (c) Mora PC, Cirri M, Allolio B, Carli F, Mura P. Enhancement of dehydroepiandrosterone solubility and bioavailability by ternary complexation with α-cyclodextrin and glycine. *J Pharm Sci* 2003;**92**(11):2177–84.
17. Ivanisevic. Physical stability studies of miscible amorphous solid dispersions. *J Pharm Sci* 2010;**99**(9):4005–12.
18. Rumondor ACF, Wilkstrom H, Van Eerdenbrugh B, Taylor LS. *AAPS PharmSciTech* 2011;**12**(4):1209–19.
19. Rumondor ACF, Ivanisevic I, Bates S, Alonzo DE, Taylor LS. Evaluation of drug-polymer miscibility in amorphous solid dispersion systems. *Pharm Res* 2009;**26**(11):2523–34.
20. Konno H, Taylor LS. Influence of different polymers on the crystallization tendency of molecularly dispersed amorphous felodipine. *J Pharm Sci* 2006;**95**(12):2692–705.
21. Newman A, Munson E. Characterizing miscibility in amorphous solid dispersions. *Amer Pharm Rev* 2012; April;92–8.
22. Vogt G. Chapter 4. Solid-state characterization of amorphous dispersions. In: Newman A, editor. *Pharmaceutical amorphous solid dispersions*. Hoboken, NJ: John Wiley and Sons; 2015.
23. Uekema K, Ikegami K, Wang Z, Joriuchi Y, Hirayama F. Inhibitory effect of 2-hydroxypropyl-β-cyclodextrin on crystal-growth of nifedipine during storage: superior dissolution and oral bioavailability compared with polyvinylpyrrolidone K-30. *J Pharm Pharmacol* 1992;**4**:73–8.
24. Newman A, Engers D, Bates S, Ivanisevic I, Kelly RC, Zografi G. Characterization of amorphous API:Polymer mixtures using X-ray powder diffraction. *J Pharm Sci* 2008;**97**(11):4840–56.
25. Ivanesevic I, Bates S, Chen P. Novel methods for the assessment of miscibility of amorphous drug—polymer dispersions. *J Pharm Sci* 2009;**98**(9):3373–86.
26. Kennedy M, Hu J, Gao P, Li L, Ali-Reynolds A, Chal B, et al. Enhanced bioavailability of a poorly soluble VR1 antagonist using an amorphous solid dispersion approach: a case study. *Mol Pharm* 2008;**5**:981–93.
27. Fox TG. Influence of diluent and copolymer composition on the glass transition temperature of a polymer system. *Bull Am Phys Soc* 1956;**1**:123–8.
28. Pham TN, Watson SA, Edwards AJ, Chavda M, Clawson JS, Strohmeier M, et al. *Mol Pharm* 2010;**7**:1667–91.
29. Andronis V, Yoshioka M, Zografi G. Effects of sorbed water on the crystallization of indomethacin from the amorphous state. *J Pharm Sci* 1997;**86**:346–51.
30. Crowley KJ, Zografi G. Water vapor absorption into amorphous hydrophobic drug'poly(vinylpyrrolidone) dispersions. *J Pharm Sci* 2002;**91**(10):2150–65.
31. Heljo VP, Nordberg A, Tenho M, Virtanen T, Jouppla K, Salonen J, et al. The effect of water plasticization on the molecular mobility and crystallization tendency of amorphous disaccharides. *Pharm Res* 2012;**29**:2684–97.
32. Taylor LS, Zografi G. Spectroscopic characterization of interactions between PVP and indomethacin in amorphous molecular dispersions. *Pharm Res* 1997;**14**:1691–8.
33. Rumondor ACF, Marsac PJ, Standford LA, Taylor LS. Phase behavior of poly(vinylpyrrolidone) containing amorphous solid dispersions in the presence of moisture. *Mol Pharm* 2009;**6**:1492–505.
34. Vogt FG, Yin H, Forcino RG, Wu L. ^{17}O solid-state NMR as a sensitive probe of hydrogen bonding in crystalline and amorphous solid forms of diflunisal. *Mol Pharm* 2013;**10**:3433–46.
35. Sinclair W, Leane M, Clarke G, Dennis A, Tobyn M, Timmins P. Physical stability and recrystallization kinetics of amorphous ibipinabant drug product by fourier Raman spectroscopy. *J Pharm Sci* 2011;**100**(11):4687–99.
36. Furyama N, Hasegawa S, Hamaura T, Yada S, Nakagami H, Yonemochi E, et al. Evaluation of solid dispersions on a molecular level by the Raman mapping technique. *Int J Pharm* 2008;**361**(1–2):12–18.
37. Aso Y, Yoshioka S, Miyazaki T, Kawanishi T, Tanaka K, Kitamura S, et al. Miscibility of nifedipine and hydrophilic polymers as measured by ^{1}H-NMR spin-lattice relaxations. *Chem Pharm Bull* 2007;**55**(8):1227–31.
38. Tatton AS, Pham RN, Vogt FG, Iuga D, Edwards AJ, Brown SP. Probing hydrogen bonding in cocrystals and amorphous dispersions using 12N-1H QC solid-state NMR. *J Pharm* 2012;**101**(2):641–63.
39. Aso, et al. *J Pharm Sci* 2006;**95**(2):318–25.
40. Aso, et al. *Chem Pharm Bull* 2009;**57**(1):61–4.
41. Patel, et al. *J Pharm Sci* 2012;**101**(2):641–63.

42. Calahan JL, Zanon RL, Alvarez-Nunez F, Munson EJ. Isothermal microcalorimetry to investigate the phase separation for amorphous solid dispersions of AMG 517 with HPMC-AS. *Mol Pharm* 2013;**10**(5):1949–57.
43. Newman AW, Reutzel-Edens SM, Zografi G. Characterization of the "hygroscopic" properties of active pharmaceutical ingredients. *J Pharm Sci* 2008;**97**:1047–59.
44. Rumondor ACF, Talor LS. Effect of polymer hygrocopicity on the phase behavior of amorphous solid dispersions in the presence of moisture. *Mol Pharm* 2010;**7**(2):477–90.
45. Yang J, Grey K, Doney J. An improved kinetics approach to describe the physical stability of amorphous solid dispersions. *Int J Pharm* 2010;**384**:24–31.
46. Newman A. X-ray powder diffraction in solid form screening and selection. *Amer Pharm Rev* 2011;Sept–Oct:44–51.
47. Engers D, Teng J, Jimenez-Novoa J, Gent P, Hossack S, Campbell C, et al. A solid-state approach to enable early development compounds: selection and animal bioavailability studies of an itraconazole amorphous solid dispersion. *J Pharm Sci* 2010;**99**(9):3901–22.
48. Dobry DE, Settell DM, Baumann JM, Ray RJ, Graham LJ, Beyerinck RA. A model-based methodology for spray-drying process development. *J Pharm Innov* 2009;**4**(3):133–42.
49. Crowley MM, Zhang F, Repka MA, Thumma S, Upadhye SB, Kumar S, et al. Pharmaceutical applications of hot-melt extrusion: part I. *Drug Dev Ind Phar* 2007;**33**:909–26.
50. Newman A, Knipp G, Zografi G. Assessing the performance of amorphous solid dispersions. *J Pharm Sci* 2012;**101**(4):1355–77.
51. Murdande SB, Pikal MJ, Shanker RM, Bogner RH. Solubility advantage of amorphous pharmaceuticals, part 3: is maximum solubility advantage experimentally attainable and sustainable? *J Pharm Sci* 2011;**100**:4349–56.
52. Matsumoto T, Zografi G. Physical properties of solid molecular dispersions of indomethacin with poly(vinylpyrrolidone) and poly(vinylpyrrolidone-co-vinyl-acetate) in relation to indomethacin crystallization. *Pharm Res* 1999;**16**:1722–8.
53. Lu Q, Zografi G. Phase behavior of binary and ternary amorphous mixtures containing indomethacin, citric acid, and PVP. *Pharm Res* 1998;**15**(8):1202–6.
54. Shanbhag A, Rabel S, Nauka E, Casadevall G, Shivanand P, Eichenbaum G, et al. Method for screening of solid dispersion formulations of low-solubility compounds-miniaturization and automation of solvent casting and dissolution testing. *Int J Pharm* 2008;**351**:209–18.
55. Moser JD, Broyles J, Liu L, Miller E, Wang M. Enhancing bioavailability of poorly soluble drugs using spray dried solid dispersions part I. *Am Pharm Rev* 2008;Sept–Oct:68–73.Moser JD, Broyles J, Liu L, Miller E, Wang M. Enhancing bioavailability of poorly soluble drugs using spray dried solid dispersions part II. *Am Pharm Rev* 2008;Sept–Oct:94–9.
56. Keen JM, McGinity JW, Williams RO. Enhancing bioavailability through thermal processing. *Int J Pharm* 2013;**450**(1–2):185–96.
57. Forster A, Hempenstall J, Tucker I, Rades T. The potential of small-scale fusion experiments and the Gordon-Taylor equation to predict the suitability of drug/polymer blends for melt extrusion. *Drug Dev Ind Pharm* 2001;**27**(6):549–60.
58. Nagy ZK, Balogh A, Dravavolgyi G, Ferguson J, Pataki H, Vajna B, et al. Solvent-free melt electrospinning for preparation of fast dissolving drug delivery system and comparison with solvent-based electrospun and melt extruded systems. *J Pharm Sci* 2013;**102**(2):508–17.
59. Yin SX, Grosso JA. Selecting and controlling API crystal form for pharmaceutical development- strategies and processes. *Cur Opn Drug Disc Dev* 2008;**11**:771–7.
60. Taylor LS. Chapter 5. Physical stability and crystallization inhibition. In: Newman A, editor. *Pharmaceutical amorphous solid dispersions*. Hoboken, NJ: John Wiley and Sons; 2015. p. 179–217.
61. Marsac PJ, Konno H, Taylor LS. A comparison of the physical stability of amorphous felodipine and nifedipine systems. *Pharm Res* 2006;**23**(10):2306–16.
62. Miyazaki T, Yoshioka S, Aso Y, Kojima S. Ability of polyvinylpyrrolidone and polyacrylic acid to inhibit the crystallization of amorphous acetaminophen. *J Pharm Sci* 2004;**93**(11):2710–17.
63. Yoshioka M, Hancock BC, Zografi G. Inhibition of indomethacin crystallization in poly(vinylpyrrolidone). *J Pharm Sci* 1995;**84**(8):983–6.
64. Crowley KJ, Zografi G. The use of thermal methods for predicting glass-former fragility. *Thermochim Acta* 2001;**380**(2):79–93.
65. Hancock BC, Shamblin SL. Molecular mobility of amorphous pharmaceutical determined using differential scanning calorimetry. *Thermochim Acta* 2001;**380**(2):95–107.
66. Khougaz K, Clas S-D. Crystallization inhibition in solid dispersions of MK-0591 and poly(vinylpyrrolidone) polymers. *J Pharm Sci* 2000;**89**(10):1325–34.
67. Tantishaiyakul V, Kaewnopparat N, Ingkatawornsong S. Properties of solid dispersions of piroxicam in polyvinylpyrrolidone K-30. *Int J Pharm* 1996;**143**(1):59–66.
68. Craig DQM. The mechanisms of drug release from solid dispersions in water-soluble polymers. *Int J Pharm* 2002;**231**:131–44.
69. Hseih Y-L, Ilevbare GA, Van Eerdenbrugh B, Box KJ, Sanchez-Felix MV, Taylor LS. pH-Induced precipitation behavior of weakly basic compounds: determination of extent and duration of supersaturation using potentiometric titration and correlation to solid state properties. *Pharm Res* 2012;**29**(10):2738–53.
70. Ilevbare GA, Taylor LS. Liquid-liquid phase separation in highly saturated aqueous solutions of poorly water-soluble drugs: implications for solubility enhancing formulations. *Cryst Growth Des* 2013;**13**(4):1497–509.
71. Warren DB, Menameur H, Porter CJH, Pouton CW. Using polymeric precipitation inhibitors to improve the absorption of poorly water-soluble drugs: a mechanistic basis for utility. *J Drug Target* 2010;**18**(10):704–31.
72. Ilevbare GA, Liu H, Edgar KJ, Taylor LS. Maintaining supersaturation in aqueous drug solutions: impact of different polymers in induction times. *Cryst Growth Des* 2012;**13**(2):740–51.
73. Ilevbare GA, Liu H, Edgar KJ, Taylor LS. Understanding polymer properties important for crystal growth inhibition- impact of chemically diverse polymers on solution crystal growth of ritanovier. *Cryst Growth Des* 2012;**12**(6):3133–43.
74. Ilevbare GA, Liu H, Edgar KJ, Taylor LS. Inhibition of solution crystal growth of ritanovir by cellulose polymers- factors influencing polymer effectiveness. *CrystEngComm* 2012;**14**:6503–14.
75. Chauhan H, Hui-Gu C, Atef E. Correlating the behavior of polymers in solution as precipitation inhibitor to its amorphous stabilization in solid dispersions. *J Pharm Sci* 2013;**102**(6):1924–35.
76. Dai W-G, Dong LC, Li S, Pollock-Dove C, Chen J, Manksy P, et al. Parallel screening approach to identify solubility-enhancing formulations for improved bioavailability of a poorly water-soluble compound using milligram quantities of material. *Int J Pharm* 2007;**336**:1–11.
77. Dai W-G. In vitro methods to assess drug precipitation. *Int J Pharm* 2010;**393**:1–16.
78. Davis AF, Hadgraft J. Effect of supersaturation on membrane transport: 1. Hydrocortisone acetate. *Int J Pharm* 1991;**76**(1–2):1–8.
79. Theeuwes F, Gale RM, Baker RW. Transference: a comprehensive parameter governing permeation of solutes through membranes. *J Memb Sci* 1976;**1**(1):3–16.

80. Sousa LA, Reutzel-Edens SM, Stephenson GA, Taylor LS. Assessment of the amorphous "solubility" of a group of diverse drugs using new experimental and theoretical approaches. *Mol Pharm* 2015;**12**:484–95.
81. Mosharraf M, Nystrom C. Apparent solubility of drugs in partially crystalline systems. *Drug Dev Ind Pharm* 2003;**29**(6):603–22.
82. Gao P, Shi Y. Characterization of supersaturatable formulations for improved absorption of poorly soluble drugs. *AAPS J* 2012;**14**(4):703–13.
83. Murdande SB, Pikal MJ, Shanker RM, Bogner RH. Aqueous solubility of crystalline and amorphous drugs: challenges in measurement. *Pharm Dev Technol* 2011;**16**(3):187–200.
84. General Chapter <1092> The dissolution procedure: development and validation. In: *The United States Pharmacopeia 34–National Formulary 29*. Rockville, MD: United States Pharmacopeial Convention, Inc.; 2006.
85. Carino SR, Sperry DC, Hawley M. Relative bioavailability of three different solid forms of pnu-141659 as determined with the artificial stomach-duodenum model. *J Pharm Sci* 2010;**99**:3923–30.
86. Gao Y, Carr A, Spence JK, Wang WW, Turner TM, Lipari JM, et al. A pH-dilution method for estimation of biorelevant drug solubility along the gastrointestinal tract: application to physiologically based pharmacokinetic modeling. *Mol Pharm* 2010;**7**:1516–26.
87. Polster CS, Stassi F, Wu S-Y, Sperry DC. Use of artificial stomach-duodenum model for investigation of dosing fluid effect on clinical trial variability. *Mol Pharm* 2010;**7**:1533–8.
88. Breitenbach J. Melt extrusion: from process to drug delivery technology. *Eur J Pharm Biopharm* 2002;**54**(2):107–17.
89. Crowely K, Gryczke A. Chapter 10. Hot melt extrusion of amorphous solid dispersions. In: Newman A, editor. *Pharmaceutical amorphous solid dispersions*. Hoboken, NJ: John Wiley and Sons; 2015. p. 341–63.
90. Dobry DE, Settell DM, Bauman JM. Chapter 9. Spray drying and scale-up. In: Newman A, editor. *Pharmaceutical amorphous solid dispersions*. Hoboken, NJ: John Wiley and Sons; 2015. p. 315–40.
91. Gilis PMV, DeConde VFV, Vandercruys RPG. 1997. US Patent 5,633,015.
92. Intellence Package Insert. http://www.intelence.com/shared/product/intelence/prescribing-information.pdf. Accessed October 3, 2015.
93. "Stability Testing of New Drug Substances and Products," ICH, Q1A (R2), 2005.
94. Shah N, Iyer RM, Mair H-J, Choi DS, Tian H, Diodone R, et al. Improved human bioavailability of vemurafenib, a practically insoluble drug, using an amorphous polymer-stabilized solid dispersion prepared by a solvent-controlled coprecipitate process. *J Pharm Sci* 2013;**102**:967–81.
95. Connelly PR, Quinn BP, Johnston S, Bransford P, Mudnuri P, Peresypkin A, et al. Chapter 7. Translational development of amorphous dispersions dispersions. In: Newman A, editor. *Pharmaceutical amorphous solid dispersions*. Hoboken, NJ: John Wiley and Sons; 2015.
96. Sanghvi T, Katstra J, Quinn BP, Thomas H, Hurter P. Chapter 11. Formulation development of amorphous dispersions. In: Newman A, editor. *Pharmaceutical amorphous solid dispersions*. Hoboken, NJ: John Wiley and Sons; 2015.
97. Bauer J, Spanton S, Henry R, Quick J, Dziki W, Porter W, et al. Ritonavir: an extraordinary example of conformational polymorphism. *Pharm Res* 2001;**18**:859–66.
98. Chemburkar SR, Bauer J, Deming K, Spiwek H, Patel K, Morris J, et al. Dealing with the impact of ritonavir polymorphs on late stages of bulk drug process development. *Org Proc Res Dev* 2000;**4**:413–17.
99. Law D, Krill SL, Schmitt EA, Fort JJ, Qiu Y, Wang W, et al. Physicochemical considerations in the preparation of amorphous ritonavir-poly(ethylene glycol) 8000 solid dispersions. *J Pharm Sci* 2001;**90**:1015–25.
100. Law D, Schmitt EA, March KC, Everitt EA, Wan W, Fort JJ, et al. Ritonavir-PEG 8000 amorphous solid dispersions: in vitro and in vivo evaluations. *J Pharm Sci* 2004;**93**:563–70.
101. Sham HL, Kempf DJ, Molla A, Marsh KC, Kumar GN, Chen C-M, et al. ABT-378, a highly potent inhibitor of the human immuno deficiency virus protease. *Antimicrob Agents Chemother* 1998;Dec:3218–24.
102. Klein CE, Chiu Y-L, Awni W, Zhu T, Heuser RS, Doan T, et al. The tablet formulation of lopinavir/ritonavir provides similar bioavailability to the soft-gelatin capsule formulation with less pharmacokinetic variability and diminished food effect. *J Acquired Immune Defic Syndrome* 2007;**44**:401–10.
103. Breitenbach J. Melt extrusion can bring new benefits to HIV therapy. *Am J Drug Deliv* 2006;**4**:61–4.
104. US Patent 20140288108. Process for the Preparation of Solid Dispersion of Lopinavir and Ritonavir. Sept 25, 2014.
105. US Patent 20090311414. Process for Producing a Solid Dispersion of an Active Ingredient. Dec 17, 2009.
106. http://www.thebody.com/confs/ias2005/pdfs/WeOa0206.pdf (accessed April 7, 2016).

CHAPTER

19

Rational Design of Oral Modified-Release Drug Delivery Systems

Y. Qiu[1] *and P.I. Lee*[2]

[1]AbbVie Inc., North Chicago, IL, United States [2]University of Toronto, Toronto, ON, Canada

19.1 INTRODUCTION

Modified release (MR) drug delivery systems are developed to modulate the apparent absorption and/or alter the site of release of drugs to achieve specific clinical objectives that cannot be attained with the conventional immediate release (IR) dosage forms. Possible therapeutic benefits of a properly designed MR dosage form include improved efficacy and reduced adverse events, increased convenience and patient compliance, optimized clinical performance, a greater selectivity of activity, or new indications. A clinically successful MR product with enhanced medical benefits also offers commercial advantages, such as product differentiation and/or line extension, maximized drug potential and market expansion, and increased cost effectiveness.[1–3] As a result, MR product development has been an important tool of product line extension and an integral part of product life cycle management strategy.

Drug release modification is a technique or approach by which the delivery pattern of a therapeutic agent is altered via engineering of physical, chemical, and/or biological components into delivery systems for achieving desired/target plasma drug levels defined by clinical pharmacology. According to the US Food and Drug Administration (FDA), drug release characteristics of time, course and/or location of a MR system are chosen to accomplish therapeutic or convenience objectives not offered by conventional dosage forms.[4] More specifically, MR solid oral dosage forms include extended release (ER) and delayed release (DR) products. A DR dosage form releases a drug (or drugs) at a time other than immediately following oral administration. An ER dosage form is formulated to make the drug available over an extended period after ingestion, thus allowing a reduction in dosing frequency compared to a drug presented as a conventional dosage form (eg, a solution or an IR dosage form). For oral applications, the term "extended release" is usually interchangeable with "sustained release," "prolonged release," or "controlled release."

The design objective for modifying oral drug release is to alter the rate of drug input (dissolution/absorption) in the intestinal lumen to achieve a predetermined plasma profile. Common modes of oral MR delivery include (1) delayed release (eg, using an enteric coating); site-specific or timed release (eg, for colonic delivery); (2) extended release (eg, zero-order, first-order, biphasic release); and (3) programmed release (eg, pulsatile, delayed extended-release). Fig. 19.1 illustrates some examples of MR delivery profiles. Significant clinical and commercial advances have been made over the last two decades in delivering therapeutic agents via a nonmonotonic pattern at predetermined time intervals as a result of an improved understanding of the relationship between clinical pharmacology, on the one hand, and bodily physiology and biological conditions, on the other. For example, many body functions and diseases follow circadian rhythm (eg, daily fluctuations of hormones, gastric secretion and emptying, bronchial asthma, myocardial infarction, angina, rheumatic disease, ulcer, and hypertension). By timing the administration of a programmed release device, therapeutic plasma concentration can be obtained at an optimal

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00019-4

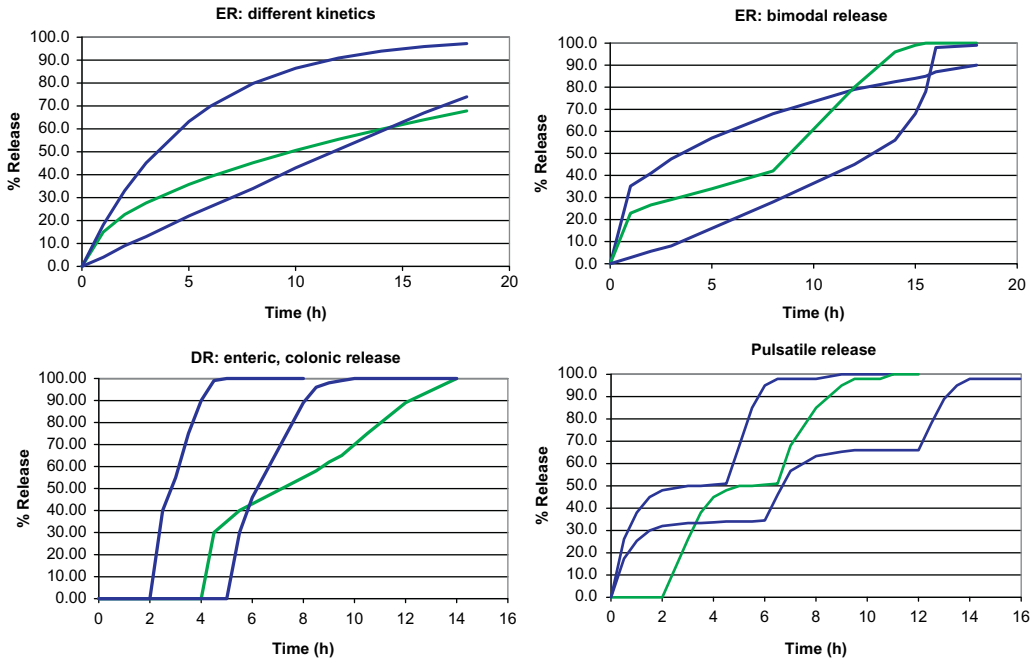


FIGURE 19.1 Examples of oral modified release profiles.

time to counter or utilizing the diurnal nature of certain diseases, such as angina, hypertension, asthmatic attacks or stiffness of arthritic patients during early morning hours, and heart attacks at night. Classic examples of utilizing MR delivery that have achieved significant clinical and commercial successes include MR products of Nifedipine, Methylphenidate, Mesalamine, Verapamil and Diltiazem, to name a few. For example, Procardia XL is a zero-order release tablet of Nifedipine that not only reduces dosing frequency from t.i.d. to once-daily but also drastically improves the efficacy-to-safety ratio. Controlling the input rate results in a gradual decrease in blood pressure without an increase in heart rate and syncope associated with t.i.d. administration. Methylphenidate is a compound indicated for attention deficit hyperactivity disorder (ADHD) without patent protection. Several new MR products of methylphenidate were introduced in the 1990s by different developers that offer clear clinical advantages over the products on the market at the time (eg, Ritalin, Ritalin SR). These products include a bimodal-extended release tablet (Concerta), a capsule (Metadate CD), and a pulsatile release capsule (Ritalin LA,). All of them were designed to produce fluctuation of blood levels over time to overcome acute tolerance associated with a constant rate of delivery, thus enabling dosing convenience of this controlled drug for school children. Asacol and Lialda are delayed release tablets designed to release mesalamine in the distal ileum and/or colon for ulcerative colitis treatment while Pentasa is an extended-release capsule that delivers the same active throughout the gastrointestinal (GI) tract. Hypertension and angina remain two of the most important risk factors for cardiovascular morbidity and mortality. MR products of cardiovascular drugs have been successfully developed to address the pharmacokinetic and circadian challenges to controlling blood pressure and angina by matching drug release to the natural circadian rhythms of the cardiovascular system; that is, chronotherapeutic drug delivery. Examples of such cardiovascular chronotherapeutics products include Adalat CC tablets that provides zero-order release followed by delayed IR of Nifedipine, as well as Covera-HS tablets and Verelan PM capsules, both of which utilize delayed extended release of verapamil. Following bedtime administration, these products provide a higher concentration of drug in the blood to protect against the early-morning increase in blood pressure and heart rate while maintaining effectual blood pressure reduction for 24 hours. Another classic example is Diltiazem hydrochloride, also a calcium antagonist for the treatment of hypertension. Three generations of dosage forms were developed with increasing clinical success and commercial competitiveness between the 1980s and 1990s. Annual revenues were approximately $260 million, $400 million, and $900 million for IR tablets (Cardizem) in 1988, twice-daily ER capsules (Cardizem SR) in 1991, and once-daily ER capsules (Cardizem CD) in 1996, respectively. Several years after patent expiry of the active ingredient, another new ER dosage form designed based on chronotherapy concept (Cardizem LA) was introduced to the market in 2003.

Many compounds are not suitable for MR delivery due to a variety of reasons, such as undesirable drug properties, large dose size, poor colonic permeability, lack of pharmacological rationale, or technical feasibility. Otherwise, solely for the purpose of convenience, one would expect very few marketed drug products with a dosing frequency of more than twice a day. Furthermore, despite many successes with developing oral MR products, as many or more failed attempts went unreported in the R&D laboratories of the pharmaceutical industry. Failures in the innovator companies can be partly attributed to the simple fact that the molecule of interest is not feasible for MR delivery. It can also be related to a lack of expertise in rational design and development of a robust MR product. For generic companies, unsuccessful attempts are mostly related to inadequate knowledge and skill in developing MR products and/or overcoming patent hurdles because the feasibility of the active has already been proven by the innovator's product. This chapter provides an overview of current MR technologies and rational approach to delivery system design. Case studies are used to illustrate how drug property and formulation influence the design and performance of oral MR delivery systems.

TABLE 19.1 Oral ER Systems Commonly Utilized in Commercial Products

	Matrix	Reservoir	Osmotic
Systems	Hydrophilic Matrix • Erosion/drug diffusion controlled • Swelling/erosion controlled Hydrophobic Matrix • Homogenous (dissolved drugs) • Heterogeneous (dispersed drugs)	Membrane Controlled Constant activity Nonconstant activity Membrane-matrix Combination	Elementary Osmotic Pump Microporous Osmotic Pump Layered Osmotic Pump (eg, Push-Pull, Push-Stick)
Common Dosage Forms	Monolithic tablet Multiparticulates (pellets orminitablets) Layered tablet Compression coated tablet	Multiparticulates (coated pellets, beads or minitablets) Monolithic tablet containing coated beads Monolithic coated tablet	Coated monolithic tablet Coated layered tablet Coated beads

19.2 ORAL MR TECHNOLOGIES AND DRUG DELIVERY SYSTEMS

Oral MR Drug delivery technology has been applied to new product development for more than 60 years. Over the past three decades, tremendous progress has been made in the advancement of theory, mathematical modeling, new rate-controlling materials, technology platforms as well as processing technologies. In particular, the emergence of high-performance polymers and aqueous based polymeric dispersion has made manufacturing of MR dosage forms more amenable to conventional means of processing. Today, oral MR technology has become more widely utilized for the in-house development of new products at both innovator and generic companies.

19.2.1 Common oral extended-release systems

Various physical and chemical approaches have been successfully applied to produce well-characterized delivery systems that extend drug release into the GI tract within the specifications of the desired release profile. Today, most proprietary and nonproprietary ER technologies are based on polymeric systems. The fundamental design principles, theoretical considerations and applications of these systems have been extensively addressed and reviewed.[1,5–8] A survey of commercial ER oral solid products indicates that most systems fall into one of three broad categories: matrix, reservoir (or membrane controlled) and osmotic systems (see Table 19.1). Drug release from these ER delivery systems generally involves one or a combination of the following mechanisms: drug diffusion (through pores of a barrier, through tortuous channels, through a viscous gel layer or through the interstitial space between polymer chains), system swelling (followed by diffusion and/or erosion and dissolution), or osmotic pressure-induced release (osmotic pressure buildup forced drug solution, suspension or wetted mass flowing out of the system). Each type of system has its advantages and shortcomings with respect to the performance, applicability, manufacture, control, development time and cost, among others.

19.2.1.1 Matrix systems

In a matrix system, the drug substance is homogeneously mixed with the rate-controlling material(s) and other inactive ingredients to form a monolithic crystalline, amorphous, or in rare cases, molecular dispersion matrix. Drug release occurs by drug

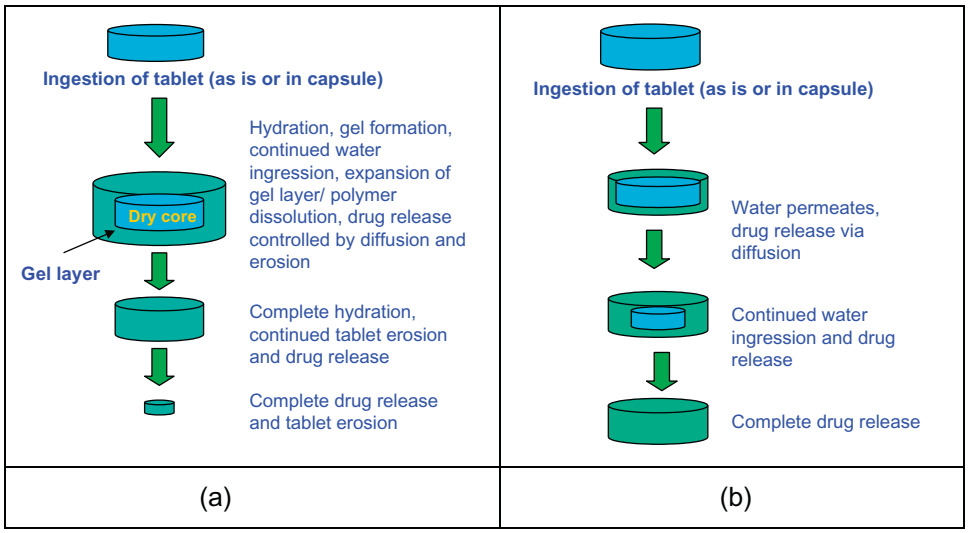


FIGURE 19.2 (a) Hydrophilic and (b) hydrophobic matrix systems and corresponding drug release process.

diffusion from and/or erosion of the matrix system. Based on the characteristics of the rate-controlling material, the matrix system can be divided into (a) hydrophilic and (b) hydrophobic systems, as shown in Fig. 19.2. For practical purposes, the former refers to a matrix system in which the rate-controlling materials are water-soluble and/or swellable, while the latter consists of a water-insoluble inert matrix with minimum swelling (ie, rigid matrix). Matrix systems consisting of a mixture of hydrophilic and hydrophobic rate-controlling materials belong to one or the other of the two systems depending on the dominant mechanism of drug release control. For instance, a matrix containing both types of polymers can be considered a hydrophilic matrix if the release control function and release kinetics (mechanism) remain essentially unchanged when the insoluble hydrophobic polymer is substituted with a conventional insoluble excipient (eg, dicalcium phosphate) that does not possess a rate-controlling property. However, at a high percentage, a nonswelling insoluble excipient may disrupt or destroy the ER properties of the matrix due to nonuniformity in the gel layer. The same holds true for a hydrophobic system if the hydrophilic polymer is substituted with a soluble excipient (eg, lactose). Additionally, the manufacturing method can also influence the primary release-control mechanism because the internal structure of a matrix can differ depending on whether the insoluble rate-controlling polymer(s) undergo solid phase transition (eg, dissolving, melting) during processing.

The matrix system has been most widely utilized to provide extended delivery of drug substances because of its effectiveness and capability of accommodating both low and high loading of drugs having a wide range of physical and chemical properties. From a product development point of view, it is cost-effective and easy to scale-up and manufacture. Additionally, this type of system is also suitable for in-house development since it is usually manufactured using conventional processes and equipment. However, the release characteristics of a matrix (eg, kinetics and pH-dependency) are usually determined by the properties of the drug substance and the rate-controlling polymer. To alter release profiles or to achieve unique release patterns (eg, biphasic or delayed ER), a more complex design and process, such as a layered or compression coated tablet or layered beads, would sometimes be required. Furthermore, a matrix system may lack flexibility in offering multiple strengths that are usually required for Phase 1–3 clinical studies in developing a new chemical entity because compositionally proportional dosage forms of different strengths may not provide the same proportional release rate in vitro and/or in vivo.[9] Thus, additional resources and development time are often required for new dosage strengths except for low dose drugs (eg, drug loading <5%).

19.2.1.1.1 Hydrophobic matrix systems

The hydrophobic matrix system was the earliest oral extended release platform for medicinal use. In fact, its prototypes can be traced back to the 2nd century BC and the 4th century AD, respectively, when animal fats and wax pills were used to prolong the medicinal effects of Chinese medicines.[10,11] For example, medical practitioners were instructed to "use wax pills for their resistance to dissolve thereby achieving the effect gradually and slowly."[12,13] In modern medicine, the matrix

technology has been successfully applied to commercial products for many decades. For example, Premarin tablets, one of the classic examples, has been on the market since 1942. In a hydrophobic inert matrix system, the drug is uniformly dissolved or dispersed throughout a matrix that involves an essentially negligible increase of the device surface area or change in dimension. For a homogeneous monolithic planar matrix system containing dispersed drug, the release behavior can be described by the Higuchi equation, which is a remarkably simple expression derived without solving the complex governing differential equations[14]:

$$M_t = [DC_s(2A - C_s)t]^{1/2} \tag{19.1}$$

M_t is the drug released per unit area at time t, A is the drug loading per unit volume, C_s is drug solubility, and D is the diffusion coefficient in the matrix phase. Eq. (19.1) was derived based on the assumptions that: (1) the initial drug loading in the matrix is greater than the drug solubility; (2) a pseudo-steady state exists (equivalent to the assumption of a linear drug concentration profile in the matrix); (3) the drug dissolution rate at the dispersed drug front is rapid; (4) a semiinfinite geometry applies and the area of the matrix is large that the edge effect can be neglected; (5) the diffusion coefficient is constant; (6) perfect sink conditions exist in the external media; (7) the distance between suspended drug particles is much smaller than the matrix thickness; and (8) no swelling or erosion but only the diffusion process occurs. In the case of $A \gg C_s$, Eq. (19.1) reduces to:

$$M_t = [2DAC_s t]^{1/2} \tag{19.2}$$

Thus, the amount of drug released is proportional to the square root of time, A, D, and C_s. Higuchi also extended this pseudo-steady state analysis to the release of suspended drug from a spherical matrix system.[15]

Drug release from a porous monolithic matrix system involves the simultaneous penetration of surrounding liquid, dissolution of the drug, and leaching out of the drug through interstitial channels or pores. The volume and length of the openings in the matrix must be accounted for in the diffusion equation, leading to a second form of the planar Higuchi equation[15]:

$$M_t = \left[\varepsilon C_s(2A - \varepsilon C_s)\frac{Da}{\tau}t\right]^{1/2} \tag{19.3}$$

where ε and τ are the porosity and tortuosity of the matrix, respectively, and D_a is the drug diffusion coefficient in the release medium. Tortuosity is introduced to account for an increase in diffusion path length due to branching and bending of the pores. Similarly, Eq. (19.9) can be derived based on pseudo-steady-state approximation ($A \gg C_s$):

$$M_t = \left[2D_a AC_s\frac{\varepsilon}{\tau}t\right]^{1/2} \tag{19.4}$$

The porosity, ε, in Eqs. (19.3) and (19.4), is the fraction of matrix that exists as pores or channels into which the surrounding liquid can ingress. It is the total porosity of the matrix after the drug has been extracted. The total porosity consists of the initial porosity, ε_a, due to air, or void space in the matrix before the leaching process begins, and the porosity created by extracting the drug, ε_d, and the water soluble excipients, ε_{ex}[16,17]:

$$\varepsilon = \varepsilon_a + \varepsilon_d + \varepsilon_{ex} = \varepsilon_a + \frac{A}{\rho} + \frac{A_{ex}}{\rho_{ex}} \tag{19.5}$$

ρ is the drug density, and ρ_{ex} and A_{ex} are the density and the concentration of water soluble excipient, respectively. When no water-soluble excipient is present in the matrix and initial porosity, ε_a, is smaller than the porosity, ε_d, Eq. (19.5) becomes:

$$\varepsilon \cong \varepsilon_d = \frac{A}{\rho} \tag{19.6}$$

Hence, Eq. (19.3) and (19.4) yield:

$$M_t = A\left[\left(2 - \frac{C_s}{\rho}\right)\frac{D_a C_s}{\tau\rho}t\right]^{1/2} \tag{19.7}$$

$$M_t = A\left[\frac{2D_a C_s}{\tau\rho}t\right]^{1/2} \tag{19.8}$$

Similar to Eqs. (19.1) and (19.2), a square-root-of-time release profile is expected with a porous monolith. In contrast to the homogeneous matrix system, the release from such system is proportional to drug loading, A.

It should be noted that (1) the Higuchi equation was originally derived for planar diffusion into a perfect sink; (2) the square-root-of-time relationship only fits data up to approximately 2/3 of the total release; and (3) the Higuchi equation is strictly valid only when the drug loading is in great excess of the drug solubility; that is, $A \gg C_s$. In fact, it can be shown that, at the limit of $A \to C_s$ which often occurs with drugs of high aqueous solubility, the Higuchi equation tends to fail giving an error of 11.38% as compared to the exact solution due to the failure of the pseudo-steady state assumption at this limit.[18] To provide a more practical and accurate equation that applies to all A/C_s values, Lee presented an approximate analytical solution to this moving boundary problem without making the

pseudo-steady state assumption which has been shown to be uniformly valid for all A/C_s values:[18,19]

$$M_t = \frac{1+H}{\sqrt{3H}}\left[C_s\sqrt{D\,t}\right] \tag{19.9}$$

where

$$H = 5\left(\frac{A}{C_s}\right) - 4 + \sqrt{\left(\frac{A}{C_s}\right)^2 - 1} \tag{19.10}$$

Similar to the Higuchi equation, Eq. (19.9) also predicts a square root of time dependence for the amount of drug release per unit area but with a different dependence on A and C_s (compare with Eqs. 19.1 and 19.2). As shown by Lee, when Eq. (19.9) is applied to the analysis of drug release from planar matrix systems, the deviations from the exact results are consistently one order of magnitude smaller than that of the Higuchi equation, particularly in the case of low A/C_s values or drug loading approaching the drug solubility ($A \rightarrow C_s$). Lee also provided an approximate analytical solution to the release of suspended drug from a spherical matrix system without making the pseudo-steady state assumption.[19] To extend to porous monolithic matrix systems, one can incorporate similar approaches of Eqs. (19.3)–(19.8) and modify the working equations (eg, Eq. 19.9) accordingly, as done by Foster and Parrott.[20] It should be emphasized that an initial deviation (or a delay) from the square root of time dependence may arise due to (1) the existence of a diffusion boundary layer at the matrix surface due to insufficient mixing; (2) the drug dissolution rate at the dispersed drug front becomes slow and rate limiting especially for low solubility drugs; (3) some limited swelling of the matrix occurs; or (4) a combination of the above.[18]

A simple semiempirical exponential relation was introduced by Ritger and Peppas in 1987 to describe the general release behavior from hydrophobic non-swellable matrices in the form of slabs, spheres, and cylinders.[21]

$$Q = \frac{M_t}{M_\infty} = kt^n \tag{19.11}$$

Q is the fractional release, M_t is the drug released per unit area at time t, M_∞ is the total drug release per unit area, k is a constant, and n is the diffusional exponent. In the case of pure Fickian diffusional release, the exponent n has a limiting value of 0.50 for thin slabs, 0.45 (range of 0.43–0.50 depending on the aspect ratio) for cylindrical samples, and 0.43 for spherical samples. For $n = 1$, zero-order drug release exists and for $1 > n > 0.5$, non-Fickian diffusion behavior is observed. In the 1990s, Wu and Zhou analyzed the same mathematical problem of dispersed drug release from matrix systems using the finite element method and showed the dependence of release kinetics on the initial solute loading, the external volume, and the boundary-layer thickness. Their models involved a moving boundary problem (or Stefan-type problem) and were solved numerically to describe the entire process of diffusional release for slabs, spheres, and cylinders without the need to make any pseudo steady-state assumption.[22–25] More recently, Yonezawa et al. published a model that can be used to describe the entire release process of a heterogeneous hydrophobic matrix system made from physical mixture of wax, soluble excipients, and the actives by compensating the deviation from the pseudo-steady-state approximation ($A \gg C_s$) at the later portion of the release profiles.[26–28] In general, hydrophobic matrix systems are not suitable for insoluble drugs because the concentration gradient is too low to allow complete drug release within a reasonable time frame; that is, GI transit time.

19.2.1.1.2 Hydrophilic matrix systems

Hydrophilic matrix systems are utilized in over 75% of the extended release products on the market today[29] because they are applicable across a broad range of oral products containing high and low loading of drugs with low, medium, or high solubility, and can be manufactured using conventional equipment and processing methods. Hydrophilic matrix systems are polymer based drug delivery systems designed to control the rate of drug release during the passage of the matrix through the GI tract.[30] During GI transit, the matrices are reduced in size through surface erosion and dissolution, which reduces the probability of expulsion of an exhausted "ghost" tablets usually seen with osmotic pump, hydrophobic matrices or reservoir system. Drug release from hydrophilic matrices involves two competing mechanisms: Fickian diffusional release and matrix erosional release. The primary rate-controlling materials are polymers that hydrate and swell rapidly in an aqueous medium and form a gel layer on the surface of the system. Diffusion across the viscous gel layer is not the only drug release pathway as the erosion of the matrix following polymer swelling also contributes to the overall release. The relative contribution of each component to total release is primarily dependent upon the properties of a given drug and matrix composition. For instance, the release of a sparingly soluble drug from hydrophilic matrices involves the simultaneous ingression of water and desorption of drug via a swelling-controlled diffusion mechanism. As water penetrates into a glassy matrix and lowers the polymer glass transition temperature, the polymer swells, forms a gel layer, slowly disentangles and eventually dissolves, releasing the undissolved drug. At the same time, the dissolved drug diffuses

through this swollen gel region into the external releasing medium. This type of diffusion with concurrent swelling and erosion generally does not follow a Fickian diffusion mechanism. The continuously changing variables that affect drug release (eg, diffusion pathlength, polymer swelling, viscosity, system dimension) make obtaining a precise mechanistic equation or model describing the release profile impossible. Over the past three decades, various models have been explored and developed toward fundamental understanding of drug release from hydrophilic matrices.[6,7,19,31–37] Among them, a semiempirical exponent equation that had its root in early studies of penetrant transport in homogeneous glassy polymers,[38] but was reintroduced in the pharmaceutical literature in the mid-1980s, has been widely used to describe drug release behavior from heterogeneous hydrophilic matrix tablet systems which has the same form as Eq. (19.11)[39,40]:

$$Q = \frac{M_t}{M_\infty} = kt^n \tag{19.11a}$$

where Q is the fraction of drug released in time t, M_t is the drug released per unit area at time t, M_∞ is the total drug release per unit area, k is the rate constant specific to the characteristics of the macromolecular network system and the drug, and n is the diffusional exponent.

This equation has been widely used in this field.[41–44] It has been shown that the value of n is indicative of the drug release mechanism. For $n = 0.5$, drug release follows a Fickian diffusion mechanism which is driven by a chemical potential gradient. For $n = 1$, drug release occurs via relaxational transport, which is controlled by the swelling phase-transition and erosion in hydrated polymers. For $1 > n > 0.5$, non-Fickian diffusion behavior is often observed as a result of contributions from both diffusion and polymer erosion. This is also termed "anomalous" release. Unfortunately, over the years, Eq. (19.11) has been misused, misinterpreted or misunderstood by many in the pharmaceutical field attempting to assign a controlling mechanism based on fitting the release data to this equation to identify an exponent, without gaining any insight into the underlying physical processes and controlling events of the systems involved.

To describe the polymer dissolution process in swellable polymers, Lee and Peppas were the first to investigate this phenomenon in the field of controlled release and, through the incorporation of a pseudo-steady state assumption; they showed that at the beginning of the dissolution process the gel layer thickness increases as a square root of time.[35,45] Lee and Peppas further extended this dissolution model to arrive at an equation of the following form describing the drug release from a polymer matrix undergoing dissolution[46]:

$$Q = k_1 t^{0.5} + k_2 t \tag{19.12}$$

where the first and second terms represents drug release due to diffusion and polymer erosion, respectively, and k_1 and k_2 are constants reflecting the relative contributions of these two processes which are functions of the polymer and drug volume fractions at the gel/drug core and gel/solution interfaces as well as the drug and solvent diffusion coefficients. This model has been adopted successfully for predicting drug release from pharmaceutical tablets during swelling and dissolution.[47,48]

A more general equation describing systems of different geometries with both diffusional and relaxational contributions to release was proposed by Peppas and Sahlin, who derived the following equation by introducing a second term into Eq. (19.11)[42]:

$$Q = k_1 t^n + k_2 t^{2n} \tag{19.13}$$

In the case of a flat geometry where surface area is fixed, the value of n would be equal to 0.5, resulting in Eq. (19.12).

Since the early 1990s, additional models have been investigated in an attempt to enhance the understanding of the matrix delivery systems.[49–52] One of the examples is the so-called "Spaghetti" model proposed to gain insight into the complex release process from hydrophilic matrix systems. This model treats polymer erosion as diffusion of polymer across an unstirred "diffusion layer" adjacent to the polymer gel layer. Thus, two competitive diffusional processes contribute to overall drug release; that is, diffusion of drug through the gel layer and diffusion of polymer across the diffusion layer. In addition to the solubility of the drug molecules that defines the diffusion component, polymer disentanglement concentration ($C_{p,dis}$) defined in Eq. (19.14) is used to gauges the contribution of polymer diffusion/dissolution for hydroxypropyl methylcellulose (HPMC)[53]:

$$(C_{p,dis})_{eq} = 0.05\left(\frac{MW_p X_p}{96000}\right)^{-0.8} \tag{19.14}$$

$$M_p \approx kt^1 \tag{19.15}$$

MW_p and X_p denote the molecular weight and weight fraction of polymer in the matrix, respectively, and M_p is polymer release at time t. Interestingly, the release profile of polymer (Eq. 19.15) resembles that for an erosion-controlled system of Eq. (19.11a). It should be noted that $C_{p,dis}$ is an intrinsic property of the polymer while $(C_{p,dis})_{eq}$ is an "equivalent" $C_{p,dis}$ of

polymer in a matrix. One may consider $C_{p,dis}$ as the equivalent "solubility" of polymer, as it defines the concentration at which a polymer detaches from a pure polymer system. Conceptually, the relative contribution of both mechanisms can be characterized by the solubility ratio of the drug, C_s to $(C_{p,dis})_{eq}$:

$$\text{If } \frac{C_s}{(C_{p,dis})_{eq}} >> 1, \quad \text{then } Q_t = kt^{0.5}$$

(Release is controlled by drug diffusion)

$$\text{If } \frac{C_s}{(C_{p,dis})_{eq}} << 1, \quad \text{then } Q_t = kt^{1}$$

(Release is controlled by polymer erosion)

Thus, drug release profiles from a hydrophilic matrix vary significantly with formulation design and solubility of the drug. For insoluble compounds, the zero-order release is readily attainable as C_s values are lower than $(C_{p,dis})_{eq}$. For soluble drugs, release kinetics typically follows the square-root-of-time relationship. To achieve zero-order release, one would need to increase $(C_{p,dis})_{eq}$ in the matrix to lower the ratio of C_s to $(C_{p,dis})_{eq}$, for example, by using polymers of low MW_p or decreasing X_p according to Eq. (19.4). However, the practical feasibility of this approach is dependent on the ratio, limited choice of polymer, and the minimum MW_p and/or X_p that is often required to form a robust matrix for extended release.

The drug release from hydrophilic polymer matrix systems involves (1) diffusion of dissolved drug through the gel layer; and (2) release of dissolved drug or drug particles through matrix erosion or dissolution of the polymer gel layer. As the aqueous medium penetrates a hydrophilic matrix tablet, the initial rate of hydration of the hydrophilic polymer such as HPMC would be rapid which immediately forms a gel layer around the matrix tablet and eliminates any pre-existing pores near the tablet surface. Additional uptake of the aqueous medium into the matrix system takes place by diffusion through the gel layer. Thus, the drug release from such hydrophilic matrix tablet occurs via diffusion through this expanding gel layer and is typically represented by a square root of time dependence in the cumulative drug release with some early deviations from linearity (or delays) due to the initial matrix hydration and swelling. The thickness of this swelling gel layer increases with time initially until the outer gel surface reaches a critical disentanglement concentration and erodes. Once this occurs, the drug release typically shows a positive deviation from linearity in the square root of time plot due to the contribution from polymer erosion.[33,39] Very often, the gel layer development may reach a steady state (or a constant gel thickness) as a result of synchronization of the moving solvent and erosion fronts; the gel thickness eventually decreases as the dry core disappears and the gel layer gradually depletes due to polymer dissolution as depicted in Fig. 19.3.[19,54]

Although the drug release profiles from hydrophilic matrix systems have often been associated with two limiting mechanisms as described above including purely diffusion controlled for soluble drugs (square-root-of-time kinetics) and purely polymer erosion controlled for insoluble drugs (zero-order release), this distinction appears to be overly simplified. First of all, for poorly soluble drugs where the drug dissolution rate is slower than the polymer erosion rate, it is reasonable that the polymer erosion rate will dictate the release of the drug particles, and the released drug particles will undergo further dissolution. On the other hand, while it is true that reasonably soluble drugs will be released mostly by diffusion across the gel layer because the low drug concentration at the eroding gel surface contributes very little to the total drug release, the polymer erosion process, however, does affect the gel thickness (the faster the polymer erosion, the smaller the gel layer thickness) and thereby directly influencing the diffusional path length for the release of the dissolved drug. As described by Eq. (19.14), higher molecular weight polymers have lower disentanglement concentrations while lower molecular weight polymers have higher disentanglement concentrations. In this case, reducing the molecular weight of the polymer makes the matrix erode faster (and the gel layer thinner) because of the increase of the disentanglement concentration.

As mentioned above the development of the gel layer thickness and the resulting drug release due to diffusion and polymer erosion was first investigated by Lee and Peppas[35,45,46] employing a pseudo-steady state assumption, however, it is important to note that their final working equations are only useful in dealing with systems containing dissolved drug since the

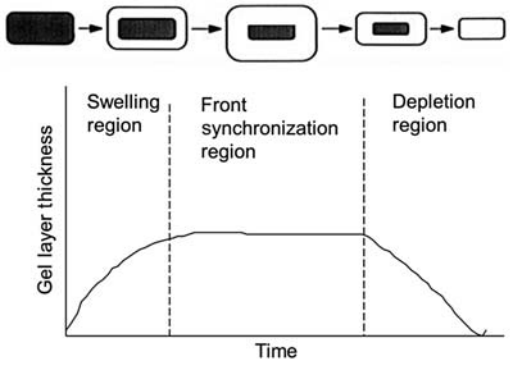


FIGURE 19.3 Dynamic gel thickness development in a hydrophilic matrix tablet.

dispersed drug situation was not analyzed. On the other hand, the problem of dispersed drug release from such a hydrophilic polymer matrix system is similar to the problem of dispersed drug release from an erodible polymer matrix investigated by Lee[19] without making the pseudo-steady state assumption. Although Lee's work does not explicitly deal with the initial swelling phase, the moving boundary problem investigated does include the release of dispersed drug due to both drug diffusion and polymer erosion, identical to the physical situation of the present hydrophilic polymer matrix system beyond the initial swelling phase. Therefore, Lee's approximate analytical solutions provide a reasonable physical picture of dispersed drug release from such a hydrophilic matrix system, in particular, they describe how the drug release is affected by the dispersed drug loading, drug solubility, polymer erosion rate constant, and drug diffusion coefficient.

More specifically, Lee's mathematical analysis[19] considers the competition between the diffusion and erosion processes in a matrix involving two moving boundaries as a function of two key parameters: drug loading to solubility ratio (A/C_s) and the relative magnitude of erosion versus diffusion (Ba/D), where a is the half thickness of the planar matrix, and B is the surface erosion rate constant. For the gel layer thickness development in a dispersed polymer matrix with $A/C_s = 5$, it can be shown in Fig. 19.4 that Lee's results predict that the faster the polymer erosion relative to diffusion (eg, $Ba/D = 10$ as compared with $Ba/D = 1$), the sooner it reaches a constant gel thickness due to front synchronization and the smaller the resulting gel thickness.

As to the effect of drug loading on the resulting drug release kinetics, Lee's results show that at a fixed relative polymer erosion to diffusion (eg, $Ba/D = 1$) the erosion characteristics (linear release region) becomes more dominant as the drug loading to solubility ratio (A/C_s) increases, whereas the early portion of the release profile remains dominated by the diffusion mechanism as illustrated in Fig. 19.5.

With respect to the effect of polymer erosion rate on drug release, Lee's results further illustrate that when the drug loading is much higher than the drug solubility (eg, $A/C_s = 5$), the polymer erosion characteristics dominate the drug release profile with increasing Ba/D as shown in Fig. 19.6. As expected, the erosion effect diminishes as the drug loading is at or below the drug solubility (ie, $A/C_s = 1$) (results not shown).

From Figs. 19.5 and 19.6, it is clear how the interplay of these parameters: drug loading to solubility ratio (A/C_s) and the relative magnitude of erosion versus diffusion (Ba/D) would affect the resulting drug release profiles. These are useful practical and

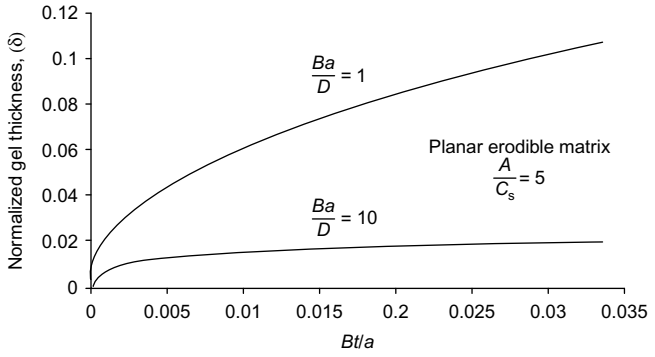


FIGURE 19.4 Effect of the magnitude of polymer erosion relative to diffusion on the gel thickness development in a dispersed erodible matrix system at $A/C_s = 5$.

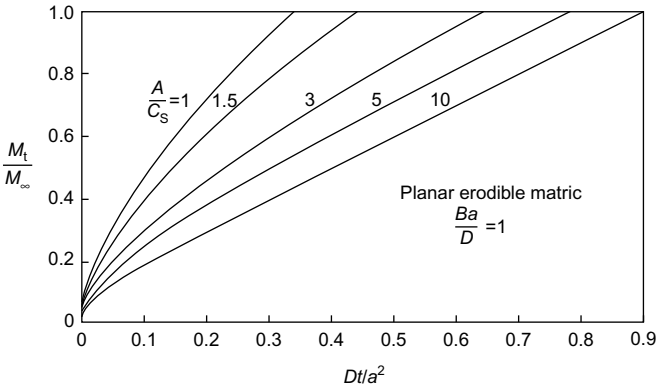


FIGURE 19.5 Effect of drug loading to solubility ratio (A/C_s) on drug release from a dispersed erodible matrix system at $Ba/D = 1$. *Figure adapted from Lee, P. Diffusional release of a solute from a polymeric matrix—approximate analytical solutions. J Membr Sci 1980;7:255–275 (reproduced with permission from the Journal of Membrane Science, Copyright Elsevier, 1980).*

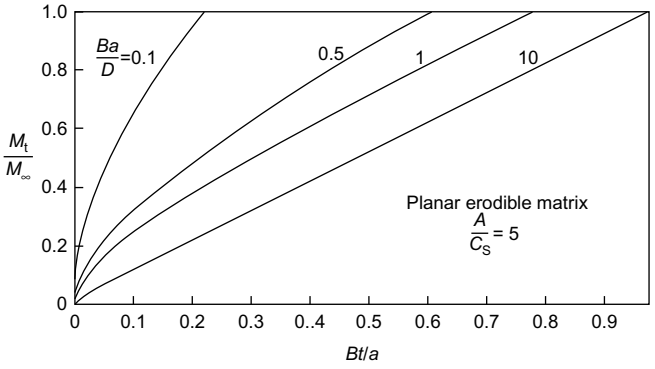


FIGURE 19.6 Effect of polymer erosion rate on drug release from a dispersed erodible matrix system at $A/C_s = 5$. *Figure adapted from Lee, P. Diffusional release of a solute from a polymeric matrix—approximate analytical solutions. J Membr Sci 1980;7:255–275 (reproduced with permission from the Journal of Membrane Science, Copyright Elsevier, 1980).*

phenomenological parameters which can be directly related to intrinsic variables such as the polymer disentanglement concentration, polymer *MW*, among others. as described in this chapter. It should be noted that although these results were obtained originally for the planar erodible matrix, the predicted general trends apply equally well to matrix tablets based on hydrophilic swellable and erodible polymers. More complex models describing the drug release from swellable and erodible matrix tablets (eg, HPMC-based) including the geometry effect have been investigated,[55] however those normally require numerical solutions every time a parameter is changed and therefore, is more cumbersome to use than approximate analytical solutions described above.

19.2.1.1.3 Modulation of drug release profile

Historically, constant-rate delivery has been one of the primary target profiles of extended-release systems for maximized coverage and minimized fluctuation in plasma concentrations, especially for drugs with a narrow therapeutic index. In a diffusion-controlled matrix system, the active agent has a progressively longer distance to travel and smaller releasing surface as the diffusion front moves inwardly, resulting in decreasing release rate over time; that is, nonlinear release characteristics. For hydrophilic matrices, the extent of drug release deviation from zero-order kinetics depends on the ratio of C_s to $(C_{p,dis})_{eq}$and A/Cs. Drugs with relatively high C_s typically exhibit a diffusion-controlled release profile. To overcome the inherently nonzero-order kinetics due to decreasing surface area and increasing diffusional path length associated with the diffusion-controlled matrix systems, considerable efforts have been and continue to be expended in modifying the delivery system. Over the past several decades, many creative designs have been reported that effectively alter the inherent nonlinear release behavior.[56–69] For example, nonuniform drug loading was used to offset the decrease in release rate by increasing the diffusional driving force over time. Geometry factors including cone shape, biconcave, donut shape, hemisphere with cavity, core-in-cup, among others, were utilized to compensate the decreasing release rate by increasing drug release surface over time. Control of matrix erosion, surface area and swelling, and matrix-membrane combinations have also been shown to be effective in providing zero-order or near zero-order release kinetics. However, many of these designs are more difficult to manufacture or impractical for commercialization.

With the growing need for optimizing clinical therapy and improved understanding of the pharmacokinetics-pharmacodynamics (PK-PD) relationship, additional delivery profiles other than zero-order kinetics are often desired, such as the profiles illustrated in Fig. 19.1. These targeted extended delivery patterns vary with drugs depending on their respective clinical pharmacology. To produce matrix systems having some of these unique drug release patterns or to overcome the inherent limitations such as pH-dependency of drug release, modifications to both hydrophilic and hydrophobic matrices have been investigated. Many of them have been successfully applied to commercial products. Among them are functionally coated matrix, layered tablet, compression coating, combined use of multiple polymers, functional excipients or geometry. For example, TIMERx and Geminex (bilayer) matrix systems utilize the synergistic interaction of xantham gum and locus bean gum in an aqueous environment to form a strong gel with a slowly-eroding core. Matrix systems containing both hydrophilic and hydrophobic rate-controlling materials also offer certain advantages.[70] Below are examples in specific areas in which significant progress has been made over the years in improving matrix delivery system. Many of the functional enhancements are achieved via utilization of interactions between polymer and drug, polymer and polymer, drug and excipients or polymer-excipients.

19.2.1.1.3.1 pH-independent drug release The rate of drug release of weakly basic or acidic drugs from matrix or reservoir systems usually differs with the environmental pH change due to the pH-dependent solubility (driving force) of the drugs and/or rate-controlling polymers. In principle, pH-independent release is desirable for a more consistent in vivo delivery because pH values in the GI tract vary with location and food intake. To achieve pH-independent drug release, pH modifying excipients are commonly incorporated in the formulation to provide constant local pH inside a dosage form.[71–75] However, the effectiveness of this approach is highly dependent on the quantity and properties of both API and the pH-modifiers, including pK_a, solubility, and molecular weight. For example, to ensure effectiveness and adequate duration of pH modulation for a weak base, an organic acid with lower pK_a and solubility is preferred because it provides increased acidity and longer residence time due to slower release. Use of highly soluble pH-modifiers can be less effective since these small molecules often leach out of the matrix system at a faster rate than that of the active, resulting in only transitory pH modification of the gel layer.[63] An example of such scenario would include an attempt to use a very soluble acid in an erosion-controlled hydrophilic matrix system or in a reservoir system requiring prolonged in vivo delivery (eg, 8–20 hours). Additionally, high concentration of the exceedingly soluble acid, such as citric or tartaric acid, in the matrix can create high local

osmotic pressure and/or ionic strength that would enhance initial water penetration, alter properties of the rate-controlling polymer, interfere with barrier properties of the gel layer and result in a change in hydration/gelling behaviors, release mechanism or overall drug release.[76–78] Some of the mechanisms involved include lowering of sol-gel transition temperature of the polymer, the suppression of water-polymer interaction that reduces polymer particle swelling and gel layer coalescence, hence disrupting the establishment of the diffusion barrier properties of the gel layer, and an increased water penetration. There was also some evidence for an increased gel layer fragility and susceptibility to erosion.[76] Furthermore, other properties of the pH-modifier that may influence processing and stability should also be taken into consideration. For instance, compression and blending of the matrix tablets may be affected by the bulk and compaction properties of the pH-modifier when a relatively high percentage is incorporated in the formulation. A significant change of dissolution on storage can occur as a result of gradually altered polymer properties including degradation, hydration/gelling behavior due to long-term exposure to acidic and/or highly osmotic and ionic microenvironment created by high hygroscopicity of the pH-modifier. For example, the viscosity of HPMC solution is known to be stable within pH of 3–11 but tends to decrease on storage under acidic or alkaline conditions.[79] Certain soluble acids, such as citric acid, can create an acidic environment with pH ≤ 2 at a concentration $\geq 2\%$, compromising HPMC performance. Wells et al. reported a drastic increase in dissolution rate of the original Wellbutrin SR tablets containing HPMC and cysteine hydrochloride, an acidic additive, at storage condition of 25°C/65%RH, resulting in significant reduction of product shelf-life to 9 months.[32,80] Through investigation, it was found that the accelerated dissolution could be primarily attributed to the slow acid hydrolysis of HPMC on storage. The problem was later resolved through product reformulation by increasing HPMC content to slow the initial dissolution and increase the ratio of HPMC to acidic additive.

To reduce or minimize the impact of these potential performances and processing issues, ionic polymers such as alginate, anionic polymers containing methacrylic acid or phthalate functional groups, cationic polymers with dimethylamino ethyl methacrylate, among others, in a matrix system have also been used and shown effective in maintaining a constant local pH environment.[74,81–83] Howard et al. and Zhang et al. showed the pH-independent zero-order release of a soluble basic drug (Verapamil HCL) using alginate/HPMC[84] and Alginate/HPMC/enteric polymer,[85] respectively. Rao et al. investigated combined HPMC/cationic polymer for pH-independent release of acidic molecules.[86] The advantages of using polymer as pH-modifiers include long residence time in the matrix, generally favorable processing characteristics, avoidance of extremely acidic, basic or highly osmotic and ionic microenvironment and contribution to the release rate control by certain polymers (eg, alginate). One possible limitation of using anionic polymers as pH modifiers includes its relatively higher pK_a (4–5). Depending on the pK_a, pH-solubility profiles and required dose/drug loading of the individual API, the effectiveness of polymeric pH-modifiers may be limited for certain formulations that require stronger buffering capacity. Lastly, it is important to understand whether a selected pH-modifier may interact with either the drug molecules or rate-controlling polymers. Such interactions may result in a undesired outcome in product performance and/or processing, such as compression, slow drug release or disruption of gel structure.[74,87] In some cases, excipient-polymer interactions may be utilized to adjust drug release behaviors for reduced sensitivity to environmental variations. For example, a so-called self-correcting hydrophilic matrix having strong gels was reported to be less sensitive to both pH and stirring condition.[88,89] This self-correcting system was based on the incorporation of a high concentration of electrolytes. However, its potential advantages over a conventional matrix system in vivo remain to be confirmed, and long-term stability information is unavailable in the public domain.

In summary, the use of buffering excipients in formulating pH-independent ER systems increases the complexity that can have consequences. Besides assessing drug release performance, the influence of the pH-modifier on the formulation, the API and processing need to be understood when designing such dosage forms.

19.2.1.1.3.2 Solubility enhancement Developing ER formulations of a poorly soluble drug can sometimes be challenging. Conceptually, certain insoluble drugs may exhibit a "natural" ER behavior when a right balance between particle size distribution and intrinsic dissolution rate is achieved such that the slow particle dissolution in vivo takes place within the drug particle's transit time through the absorption window in the GI tract. However, using formulation technology to control the release rate is more desirable for ensuring delivery control, consistency, and completion within GI residence time. Thus, incorporating more soluble or faster dissolving form of API in a matrix system is sometimes required for this purpose as well as for maximizing absorption of the drug released in the lower bowl where a limited amount of water is available for dissolution. Increased solubility and/or dissolution rate can be obtained through use of the

amorphous drug, a soluble salt, a complex or API nanoparticles. Rao et al. showed that both enhanced and pH-independent release of insoluble drugs were achieved using cyclodextrin as a complexation agent in an HPMC-based matrix.[86] Because of the slow eroding nature of the matrix system, complexation can take place in situ. Thus, it does not require a prior formation of the complex using a solution-mediated process. Tanaka et al. reported an eroding hydrophobic matrix system containing an amorphous solid dispersion of nilvadipine with HPMC.[90] Supersaturation was achieved without any recrystallization during dissolution, likely a result of crystallization inhibition by HPMC. In addition to that, the release rate was found to be independent of pH and agitation.

19.2.1.1.3.3 Modification of release kinetics One of the limitations in controlling the release rate of soluble drugs using a hydrophobic or hydrophilic matrix system is the inherently nonlinear release kinetics. In addition to the use of unique geometry or nonuniform drug loading mentioned above, commercially viable designs have been utilized to achieve zero-order drug release. For instance, multilayered matrix systems were shown to be effective through control of swelling and surface[91]. Hydrophobic matrices were press-coated with hydrophilic and/or hydrophobic barrier layer(s) to provide delayed drug release from the barrier surface to compensate for the decreasing release rate over time.[92] The same type of design can also be used to control effectively the release rate of highly soluble drugs that are difficult to slow down using monolithic hydrophilic or hydrophobic matrices.[93] Successful use of synergies of polymers for release control of soluble compounds was proven in vitro and in vivo with the TIMERx delivery matrix system. In studying a polymer complex between methylcellulose (MC) and carboxyvinylpolymer (CP) on the release characteristics of phenacetin (PHE), Ozeki et al. showed that the release profile of PHE from a 20% solid dispersion granules in MC−CP complex can be modulated via altering the MC/CP ratio and the molecular weight of MC. Contributions of diffusion and polymer erosion to PHE release increased as the molecular weight of the MC increased.[94] The synergistic effect of Starch 1500 in an HPMC-based matrix was claimed for a soluble drug molecule in vitro.[95] However, the underlying mechanism of this observation and in vivo performance require further investigation. While the desired outcome of release modification may be achieved by utilizing interactions between polymer and drug or excipients, it is always beneficial to investigate the underlying mechanism of the interactions based on both theoretical consideration and advanced analytical techniques in order to evaluate the practical utility, application, control and limitations.[96,97]

19.2.1.2 Reservoir polymeric systems

A typical reservoir system consists of a core containing a solid drug or highly concentrated drug solution surrounded by a film or membrane of a rate-controlling material. In this design, the only structure effectively limiting the release of the drug is the polymer layer surrounding the reservoir. Based on Fick's first law of diffusion, the governing one-dimensional release rate of a drug from a reservoir system at steady state is given by:

$$\frac{dM_t}{dt} = \frac{DSK}{L}\Delta C \tag{19.16}$$

M_t is the total amount of drug released at time t, D is the diffusion coefficient of the drug, S is the effective membrane or barrier surface area for drug diffusion, L is the diffusional pathlength (eg, thickness of the film), K is the partition coefficient of drug between the barrier membrane and external aqueous phases, and ΔC is the drug concentration gradient between the solubility, C_s in the reservoir and the drug concentration, C_e in the external aqueous medium. Since the membrane coating is essentially uniform in composition and thickness, for a given molecule and system composition, D, S, K, L, and ΔC are constant in Eq. (19.16) under sink conditions ($C_s \gg C_e$). Thus, the amount of drug released as a function of time can be obtained by integration:

$$M_t = \left[\frac{DSK\Delta C}{L}\right] t = kt \tag{19.17}$$

where k is the release rate constant. The driving force of such systems is the concentration gradient of active molecules between reservoir and sink. Thus, the drug release from this type of system follows apparent zero-order kinetics until ΔC is no longer constant due to complete dissolution of the solid drug in the core. Because the reservoir system relies on ΔC as the driving force for drug diffusion, applies to soluble drugs. For insoluble drugs, the values of C_s may be too low to render adequate driving force, resulting in over-attenuated and incomplete drug release.

In developing oral products based on ER reservoir technology, polymer film coated beads or tablets and microencapsulates (microparticles, pellets or minitablets) are common dosage form presentations. A polymer membrane that contains the drug also includes a hydrophilic and/or leachable additive (eg, a second polymer, surfactant or plasticizer) to give a porous device, offering a predetermined resistance to drug diffusion from the reservoir to the sink. The resistance provided is a function of film thickness and characteristics of both the film and the migrating species in a given environment. In the real-world products, the

mechanisms of drug release from the film-coated dosage forms can be categorized into (1) transport of the drug through a network of capillaries filled with dissolution media ($K = 1$); (2) transport of the drug through the homogeneous film barrier by diffusion; (3) transport of the drug through a hydrated swollen film; and (4) transport of the drug through flaws, cracks and imperfections within the coating matrix ($K = 1$).[98–100] The key factors affecting drug diffusion are polymers and pore former in the membrane coat, drug load, and solubility.[101] The preferred reservoir system normally consists of many coated units such as beads, pellets, and minitablets. In fact, most of the marketed products based on the reservoir system are multiunit dosage forms. Unlike a single-unit tablet, the number of particulates of a reservoir system is often sufficient to minimize or eliminate the impact of any coating defect associated with a limited number of units. An important feature of a multiunit system is that tailored drug release can be readily obtained by combining subunits with different release characteristics. The multiunit system is also adaptable to varying dose strengths without the need for a new formulation. This is highly desirable during clinical development program of the new drug candidates, where dose levels are frequently adjusted based on study outcome.

Similar to matrix systems, drug release from a reservoir system usually varies with pH unless the solubility of the active is pH-independent. To achieve pH-independent release for drugs with pH-dependent solubility, C_s in the core needs to remain unchanged. Success with incorporating buffering agents to maintain constant pH in the core has been reported. As with the pH-independent matrix system discussed previously, the effectiveness of this approach also depends on the loading, buffering capacity, relative quantity, solubility and molecular weight of the buffering agent relative to the drug. Finally, the osmotic pressure in the reservoir can become a significant factor in influencing release control depending on loading and solubility of both the active and certain soluble excipients. In some cases, the osmotic components may lead to the undesired coating rupture.

19.2.1.3 *Osmotic pump systems*

The osmotic pump system is similar in construction to a reservoir device but contains an osmotic agent that acts to imbibe water from the surrounding medium via a semipermeable membrane which is permeable to water but impermeable to drug. Such a device, called the elementary osmotic pump (EOP), was first described by Theeuwes and Higuchi (1975).[102] The delivery of the active agent from the device is controlled by water influx across the semipermeable membrane. The drug is forced out of an orifice in the device by the osmotic pressure generated by the device. The size of the orifice is designed to minimize solute diffusion, while preventing the build-up of a hydrostatic pressure head that has the effect of decreasing the osmotic pressure and changing the volume of the device.

In an osmotic system, the rate of water imbibition into the system in terms of volume can be expressed as:

$$\frac{dV}{dt} = \frac{Ak}{l}(\Delta\pi - \Delta P) \tag{19.18}$$

dV/dt is the rate of water flow, k is the hydraulic permeability, A is the membrane area, l is the thickness, $\Delta\pi$ is the osmotic pressure difference, and ΔP is the hydrostatic pressure difference. When the device is rigid, the volume of the device remains constant during operation. Thus, the amount of drug released at time t can be expressed by

$$\frac{dM}{dt} = \frac{dV}{dt}[S] \tag{19.19}$$

$[S]$ is the drug solubility. When the hydrostatic pressure difference is negligible and there is an excess of solid (saturated solution) in the device, Eq. (19.19) becomes

$$\frac{dM}{dt} = \frac{kA}{l}\Delta\pi[S] \tag{19.20}$$

Therefore, the drug release rate remains constant delivering a volume equal to the volume of osmotic water uptake.

In developing oral products, two types of osmotic pump systems have been commonly used, a one-chamber EOP system and a two-chamber system (eg, Push-Pull and Push-Stick). A comparison of different types of osmotic devices is provided in Table 19.2. In general, an EOP system is only feasible for molecules with a narrow range of solubility, for example, approximately 50–300 mg/mL, to achieve zero-order and complete release. To overcome such limitation, McClelland et al. investigated the conversion of release profiles from first-order to zero-order kinetics for highly soluble (>590 mg/mL) diltiazem hydrochloride that was released from controlled -porosity osmotic pump devices following first-order kinetics. By utilizing the common ion effect, the solubility was reduced in the presence of NaCl to 155 mg/mL. As a result, zero-order drug release was achieved over a 14–16-h period. To maintain an effective NaCl concentration within the drug compartment over the release duration, NaCl granules were coated with a microporous cellulose acetate butyrate 381-20 film.[103] In recent years, studies have shown more general application of controlled-porosity osmotic pump system to drugs

TABLE 19.2 Comparison of Selected Oral Osmotic Pump Systems

Osmotic system	Similarity	Difference
EOP	*Osmotic core* Semipermeable membrane controls water flow	*Single layer tablet* One small orifice Drug released as a solution
Push-Pull		*Bilayer tablet often latitudinally compressed* Osmotic agent in both layers one or more small orifices Drug released as a solution or suspension
Push-Stick		*Bilayer tablet longitudinally compressed* Osmotic agent in push layer One large orifice Drug released as a wet mass that requires subsequent disintegration and dissolution
EnSoTrol		*Single layer tablet* One or more small orifices Incorporation of wicking agent and solubility enhancer Drug released as a solution
Controlled Porosity and Single Composition Osmotic Tablet (SCOT)		*Single layer tablet* No predrilled orifice Drug released as a solution or wet mass through channels or cracks formed in situ
L-OROS		*Single softgel capsule* One small orifice Drug released as a liquid

with different properties.[104,105] For instance, Okimoto et al. described the successful application of a controlled-porosity osmotic pump tablet utilizing (SBE)7m-beta-cyclodextrin as both a solubilizer and an osmotic agent for drugs with varying physical properties, including five poorly soluble and two highly water-soluble drugs.[106] Sotthivirat et al. demonstrated that the incorporation of sulfobutylether-cyclodextrin results in the complete and sustained release of a sparingly water-soluble drug, prednisolone from a controlled porosity-osmotic pump pellet dosage form,[107] and further delineated the release mechanism from such controlled-porosity osmotic pump pellet system taking into consideration the viscosity effect due to changes in the cyclodextrin concentration.[108]

The two-chamber device was designed mainly to accommodate less soluble molecules and/or higher drug loading. It consists of a bilayer tablet compressed latitudinally or longitudinally with one push layer containing a highly swellable polymer and a drug layer (Fig. 19.7). In the GI tract, water is imbibed through the semipermeable membrane into both layers as a result of the osmotic activity gradient established across the membrane by the osmotic excipients. As both the drug and push layers hydrate, a drug suspension (eg, in Push-Pull) or semi-solid mass (eg, in Push-Stick) is formed in situ and the push layer begins to expand as a result of the hydration and swelling of the hydrophilic polymers. Delivery of the drug substance begins when the volumetric expansion of the osmotic push layer begins to "push" the active in the drug layer through the orifice, which is drilled through the semipermeable membrane on the drug layer side. Since rate control resides within the rate-controlling membrane, drug release is essentially insensitive to environmental effects, such as pH, agitation, and type of apparatus. However, this does not apply to the Push-Stick system, which contains a significantly larger orifice that is required for the delivery of the high payload. Since the 1990s, new single layer osmotic pump tablets have been designed and successfully applied to developing commercial products for poorly soluble drugs. The designs are based on the incorporation of solubility enhancement in the tablet core (eg, EnSoTrol) or in-situ creation of larger ports or channels for delivering less soluble compounds (eg, SCOT: Single Composition Osmotic Tablet).

In summary, classic osmotic pump systems offer zero-order release profile that are independent of the drug properties and release environment in most cases. However, fabrication of this type of system often requires specialized equipment and complex processes with a long cycle time. This is particularly true with the two-chamber systems, which often translates into higher cost, longer development time and larger numbers of formulation and processing variables to define and control. Additional drawbacks include (1) the solvent process required for semipermeable membrane coating, (2) sensitivity of the drug release to device dimension variation and membrane uniformity of individual units, and (3) delayed the onset of drug release.

19.2.1.4 Other extended-release systems

Additional systems that have been demonstrated to provide extended release are based on the formation of a less soluble complex or the use of ion-exchange resins,[109] among others. Polysalt flocculates were reported as an effective physicochemical approach to extending the release of aluminum sulfate through the formation of an insoluble complex with sodium

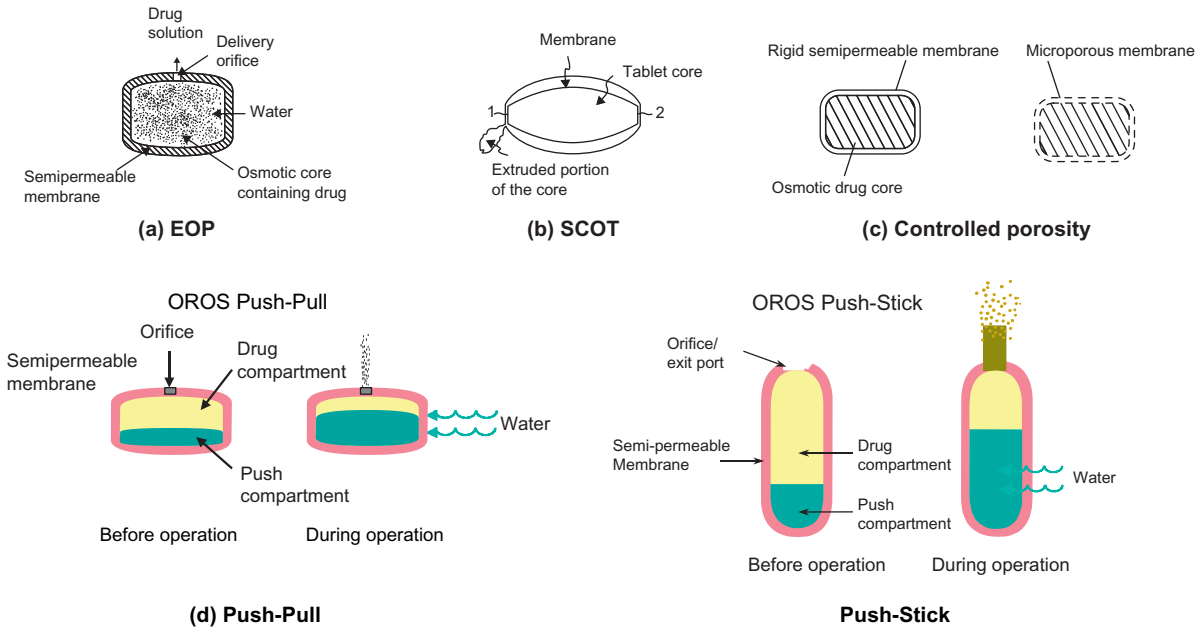


FIGURE 19.7 Schematic representation of different types of osmotic pump systems.

carboxymethylcellulose. Drug entrapment was effected by polymeric flocculation, induced by the addition of a physiologically inert electrolyte.[110] Utilization of a less soluble interaction product between diltiazem HCl and lambda carrageenan was investigated. Extended release of diltiazem from a compacted tablet was achieved via surface dissolution/erosion. The influence of complex particle size, compression force, pH of the dissolution medium, and compact dimensions on drug release was also evaluated.[111]

Synthetic polymer-based ion exchange resins were a result of advances in polymer chemistry in the middle of the 20th century. They were first developed for applications in wastewater purification and processing of fermentation products. Chaudhary and Saunders (1956) first suggested the utility of these materials for extended delivery of charged drugs.[112] Ion exchange resins are nonabsorptive polymeric insoluble particles containing basic or acidic groups in repeating positions on the polymer chain that can form ionic complexes with oppositely charged drugs.[113] The drug is bound to the resin and released by exchanging with appropriately charged ions in contact with the ion-exchange groups. The interactions are strongly governed by the pH of the medium or by the presence of competing ions. At a particular pH, one of the two entities may become neutralized depending on their respective pK_a values, thus eliminating the charge. The presence of high ionic strength buffers may reduce electrostatic interactions between the resin and drug due to a shielding/competitive binding effect. Thus, the resin may carry the drug and release the payload in a certain region of the GI tract due to a pH change or presence of competing ions. Abdekhodaie et al. developed and experimentally verified new mathematical models for drug release from ion-exchange microspheres.[114,115] The models were successfully applied to investigate the influence of important parameters pertinent to material properties and loading conditions on the kinetics, efficiency, and equilibrium of drug loading to aid the product design. Sriwongjanya and Bodmeier investigated the effect of ion exchange resins when incorporated as release modifiers into the HPMC-based hydrophilic matrix tablets.[116] Two model drugs, propranolol, and diclofenac, were studied. Drug release from matrix tablets containing drug-resin complexes was significantly slower than from the matrix tablets containing drug without resin. Pennkinetic technology developed by Pennwalt involves the use of ion exchange resins in combination with diffusion-controlling membranes to create extended-release liquid dosage forms. Examples of commercial products that utilize ion exchange resins include Tussionex Pennkinetic suspension containing hydrocodone and chlorpheniramine and Codeprex extended-release suspension of codeine polistirex/chlorpheniramine polistirex. Additional examples include Duromine containing the basic drug phentermine forming a complex with an anionic resin, and MS Contin suspension, which uses a polystyrene sulfonate resin. Medical-grade resins are available from a range of suppliers, including Rohm and Haas and Dow Chemicals. Duolite AP143 is an anionic resin, whereas Amberlite IRP64, Amberlite IRP69, and Amberlite IRP88 are cationic resins marketed by Rohm and Haas.

Drug delivery systems designed to be retained in the upper GI tract have been extensively explored and investigated for targeted and/or extended oral delivery since the 1970s. The primary objective of such systems, especially the gastroetentative devices, is to overcome regional dependency of drug absorption. For ER product, it is intended to prevent a loss of AUC due to truncated absorption of drugs with a limited window of absorption. A variety of experimental approaches that have been explored include utilizing floating (eg, low density for buoyancy) or settling mechanism (eg, high density), (2) bioadhesion, (3) altering motility with active or excipient, and (4) large size to limit gastric emptying, for example, via expansion by swelling or unfolding.[117,118] The proof of concept studies for some of these designs in humans have been reported using direct (eg, γ-scintigraphy) or indirect (PK) measures. However, none has been shown to function in a reliable and reproducible manner; that is, consistent performance under a fasting condition with acceptable intersubject and intrasubject variability of retention time. Additionally, most of the studies were carried out in a limited number of healthy subjects. For oral ER delivery, reproducibility of the gastric retention characteristics is most critical because it directly influences the bioavailability of a product. Therefore, presently none of the so-called gastroretentive technologies employed in the marketed products can be considered authentic and reliable despite the claims made by some drug delivery companies. However, it is noteworthy that an expandable system, Accordion Pill, using the design of a foldable sheet (eg, about 5 $\times$ 2.5 cm $\times$ 0.7 mm) has shown some promising in vivo results based on MRI, γ-scintigraphy, PK studies, and clinical trials. Its prolonged gastric retention with low-calorie meal has been shown in healthy volunteers.[119–121]

19.2.2 Other common oral modified-release systems

With the improved understanding of an active substance and its clinical pharmacology, various delivery profiles such as those illustrated in Fig. 19.1 are often required for enhanced or more rational and effective clinical therapy. Common examples of these nonmonotonic and multicargo delivery patterns include delayed drug delivery in the small or large intestines, pulsatile delivery, biphasic delivery (ie, IR followed ER or ER followed by IR), etc. These drug release profiles can be obtained by incorporating a range of immediate-release, delayed release, and extended release formulation approaches.

19.2.2.1 Enteric release

The enteric release is intended to delay the release of the drug (or drugs) until the dosage form has passed through the stomach.[4] The delayed liberation of orally administered drugs has been achieved through a range of formulation approaches, including single- or multiple-unit systems provided with coatings, capsule device, and osmotic pumps. The earliest physicochemical approach to delaying drug release is by applying an enteric coating to dosage forms as a barrier that controls the release location in the digestive system. Materials used for enteric coating prevent drug release in the stomach because they are typically acidic polymers that are stable/insoluble at acidic pH. The coatings dissolve rapidly at higher intestinal pH when the dosage forms reach the small intestine as shown in Fig. 19.8. Drugs such as aspirin, which have an irritant effect on the stomach, can be coated with an enteric film that will only dissolve in the small intestine. Enteric coating is also used to prevent the acidic environment of the stomach from degrading certain acid-labile medications, such as proton pump inhibitors. The common enteric dosage forms include coated tablet or capsule, coated multiparticulates in a capsule or compressed into a disintegrating tablet.

19.2.2.2 Colonic release

Targeting drug delivery to the colonic region of the GI tract has received considerable attention and is known to offer therapeutic advantages for certain drugs, such as more effective treatment of local disorders of the colon (eg, ulcerative colitis, Crohn's disease, irritable bowel syndrome and carcinoma) with reduced incidence of systemic side effects and lower dose. Colonic delivery systems are essentially delayed release dosage forms that are designed to provide either an immediate or sustained release in the large intestine. One of the better-known examples is the colonic delivery of mesalazine in the treatment of inflammatory bowel disease. More recently, scientific endeavors in this area have also been driven by the need of a portal for the systemic entry of the active substances including proteins and peptides that are unstable or poorly absorbed in the small intestine for a variety of reasons (eg, instability, metabolism). For drugs that are well absorbed throughout GI tract, sustained delivery in the colon has also been utilized in the treatment of nocturnal asthma or angina such as the previously mentioned cardiovascular chronotherapeutics products, Adalat CC and Verelan PM.

For optimal colonic delivery, the active needs to be protected from the environment of the upper GI tract before being released into the proximal colon. Based on the considerations of the unique features of the

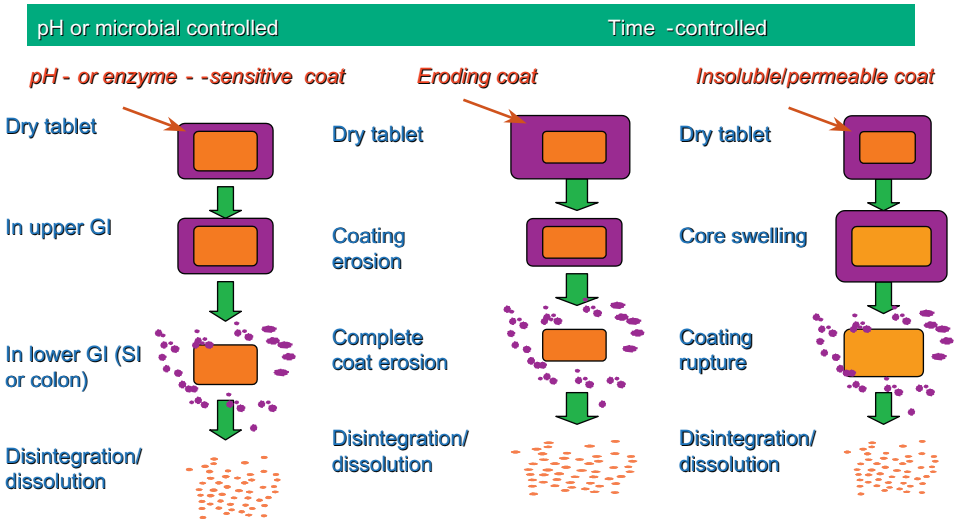


FIGURE 19.8 Examples of different types of delayed release systems and drug release process.

colonic environment: pH, transit time, pressure or microflora, various strategies and approaches have been investigated for colon-specific drug delivery. These include coated delivery systems using pH-sensitive or slow eroding polymers, swelling or controlled osmotic system for timed release and exploitation of carriers that are degraded specifically by colonic bacteria. Prodrugs have been designed to deliver 5-amino salicylic acid in the colon (eg, balsalazide, sulfasalazine, olsalazine). Certain plant polysaccharides (such as amylose, inulin, pectin and guar gum), pH-sensitive hydrogels and microbially degradable polymers especially azo cross-linked polymers have also been studied for targeting drugs to the colon. Among these approaches, coated systems that utilize the transit time and pH differential in the GI tract, as well as prodrugs that rely on colonic bacteria for release, have been utilized in commercial products.

A majority of the marketed colonic delivery products are timed release systems (see Fig. 19.8) that depend on the relatively constant transit time of 3–4 hours in the small intestine. Dosage forms such as beads or tablets are coated with enteric polymers that dissolve more slowly at intestinal pH (eg, Eudragit S). To ensure sufficiently delayed drug release of 3–4 hours following gastric emptying, the weight gain of film coating is significantly higher than that used in the enteric coating. Slowly eroding coating by compression or spraying has also been shown to be effective in timing drug release.[122] The concept of using pH alone as a trigger to release a drug in the colon is based on the pH variations along the GI tract. However, the reliability of this approach in vivo is questionable considering (1) the relative small pH difference between small and large intestines, (2) the high intersubject and intrasubject variability of intestinal pH, and (3) the impact of disease state on intestinal pH. The presence of azo reductase enzymes and glycosidase activity of the colonic microflora is the basis for designing prodrugs or polymers with a triggering mechanism for the release of drug in the colon. The release of drugs from azo bonded prodrugs or azo polymer coated dosage forms is supposed to take place after reduction and thus cleavage of the azo bonds by the azoreductase enzymes present in the colonic microflora. Similarly, natural polysaccharides that remain intact in the upper GI tract may degrade by polysaccharidases, providing site-specific drug release into the colon. To protect polysaccharides from swelling and dissolving in the stomach and small intestine, chemical cross-linking or use of a protective coat has been used. In principle, exploitation of the most distinctive property of the colon, abundant microflora, offers the most accurate site-specificity and delivery consistency in vivo compared with reliance on transit time or pH. However, applications of these modified polymers or prodrugs in commercial products have been limited primarily due to the additional resources, cost and time required for regulatory approval. Over the last two decades, the designs and evaluation of colonic drug delivery systems, as well as manufacturing considerations, have been extensively reviewed in the literature.[123–128]

19.2.2.3 *Pulsatile release*

Pulsatile delivery refers to the release of a portion of the total payload in a burst followed by periods of little/no release (lag phase) in a defined temporal pattern. In particular, oral pulsatile drug release pertains to the burst delivery of drugs following a programmed pattern from the time of oral administration. For example, Ritalin LA capsule is a pulsatile delivery system that provides IR of 50% of the total dose upon oral ingestion followed by a burst release of the remaining

drug after 4 hours. In the field of MR, these nonmonotonic and multicargo release profiles have been recognized and/or proven to offer clinical benefits in (1) optimizing chronotherapy, (2) mimicking natural patterns of endogenous secretion, and (3) providing optimal therapy for tolerance-inducing drugs where constant levels cause receptor down-regulation. Therefore, delivery systems with a pulsatile release pattern have received increasing interest in the product development.

A variety of pulsatile release systems have been investigated and successfully applied in commercial products. The fundamental system design is based on the combination of a range of formulation approaches, including single- or multiple-unit IR and delayed release systems discussed in Sections 1.2.2.1 and 1.2.2.2. The delayed release component in pulsatile delivery systems includes site-specific systems in which the drug is released at the desired site within the intestinal tract or time-controlled devices in which the drug is released after a well-defined time period. The literature reviews have provided detailed summaries of design rationale, the prominent design strategies and various single- and multiple-unit oral pulsatile delivery systems, including Pulsincap, Pulsys, and PORT Technologies.[129–134]

19.2.2.4 Bimodal release

Among nonmonotonic extended release patterns, bimodal (or biphasic) delivery profiles have been most commonly utilized to alter plasma level. The usual rationale for such designs include (1) providing rapid onset of action by adding an IR component to an extended release dosage form, (2) optimizing dosing schedules for chronotherapeutic drugs by incorporating a delayed release component in an extended release dosage form, (3) generating fluctuations of plasma levels to avoid or attenuate the development of acute tolerance due to constant exposure of drug at the receptor site, and (4) overcoming the problems associated with nonlinear pharmacokinetics, extensive first-pass metabolism, idiosyncratic pharmacokinetics or pharmacodynamics resulting in reduced bioavailability or altered drug/metabolite ratios.[135–138] Since the 1980s, many marketed products with biphasic drug release have been developed for various drugs, such as verapamil, diltiazem, nifedipine, methylphenidate, acetaminophen and zolpidem.

As with pulsatile delivery systems, biphasic release profiles are often achieved utilizing a range of formulation approaches, including single- or multiple-unit IR, delayed release and extended release systems based on the coating, matrix, and osmotic pump technologies or combinations thereof. Among common designs that have been applied are mixing systems with varying release rates, using layered tablets or multiwalled coatings, compression coated tablets with a slow eroding outer layer, combinations of delayed release coatings with osmotic pumps, among others. For example, Cardizem CD consisting of a rapid release bead and an extended release bead producing a unique "stair-step release profile."[111] Adalat CC is a compression-coated matrix tablet that provides zero-order sustained release followed by a delayed burst release. Lodotra also uses press-coated tablets (GeoClock) that provide rapid release of prednisone about 4 hours after administration at bedtime for a more effective treatment of the morning symptoms of rheumatoid arthritis.[139] Tylenol 8-hour is a bi-layer caplet designed to provide both fast and long-lasting pain relief. The first layer provides IR for fast relief and the second maintains the blood level of acetaminophen during the dosing interval. A similar design was discussed by Radebaugh et al.[140] Eichel described a delayed and sustained-release system comprising a multiwalled coated drug having an inner wall microencapsular enteric coating, a solid acid either incorporated in the enteric layer or layered over the enteric layer, and an outer microporous membrane wall for release rate control. The solid acid delays drug release by maintaining the enteric polymer in an impermeable state until the acid diffuses out of the drug or is neutralized. The multiwalled coated drug is admixed with an uncoated drug to provide biphasic release profile.[141]

19.2.3 Materials used for modifying drug release

Commonly used materials for modifying drug release from oral solid dosage forms can be categorized into long chain substituted or unsubstituted hydrocarbons and polymers. Natural or synthetic long chain hydrocarbons such as fatty acids, fatty acid esters, glyceryl esters, alcohol esters and waxes were among the earliest materials applied in modifying drug release from matrix systems. The use of waxes for prolonging the medicinal effect of herbal medicines can be traced back to the 4th century AD.[10,11] Polymers are sourced from natural products (eg, polysaccharides), chemically modified natural products (eg, cellulose ethers and esters) or synthetic in nature (eg, methacrylic ester copolymers). Today, polymers have become the dominating rate-controlling excipients in the MR arena due to their multitude of functionalities and relatively more consistent properties, especially those of synthetic or semisynthetic origin when compared with materials of natural origin. Some of the common materials that are approved for oral

administration are briefly discussed in this section based on their applications in different types of MR systems. The list is not intended to be exhaustive, but rather serves as a starting point for the interested reader.

19.2.3.1 Materials for matrix systems

The materials most widely used in preparing matrix systems include hydrophilic and hydrophobic polymers, as well as long chain hydrocarbons. Commonly available hydrophilic polymers include (1) nonionic soluble cellulose ethers available in a wide range of viscosity grades, such as HPMC (eg, Benecel K100LV, K250, K750, K1500, K4M, K15M, K35M, K100M and K200M, Methocel K100 LV, K4M, K15M, K100M; Metolose 100, 4000, 15,000, and 100,000 SR), hydroxypropyl cellulose (HPC, eg, Klucel GXF, MXF, HXF), hydroxyethyl cellulose (HEC, eg, Natrosol 250 HHX, HX, M, G) with varying degrees of substitutions and viscosity grades; (2) nonionic homopolymers of ethylene oxide, such as, poly(ethylene oxide), [H$(OCH_2CH_2)_n$OH] with a molecular weight range of 100,000 to 7,000,000 (eg, Polyox WSR N-12K, WSR N-60K, WSR-301, WSR-coagulant, WSR-303); (3) water-soluble natural gums of polysaccharides of natural origin, such as xantham gum, alginate and locus bean gum; (4) water swellable, but insoluble, high molecular weight homopolymers and copolymers of acrylic acid chemically cross-linked with polyalkenyl alcohols with varying degree of cross-linking or particle size (Carbopol 71G NF, 971P, and 974P); (5) cross-linked high amylose starch. In most cases these polymers can perform qualitatively similar functions, but possess different release-retarding capability and processing characteristics when used at the same level. They are usually available in micronized forms for facilitating the rapid formation of gelatinous barrier layer on the system surface.

Fatty acids, fatty acid esters, mono-, di-, and triglycerides of fatty acids, fatty alcohols, waxes of natural and synthetic origins with differing melting points as well as hydrophobic polymers are used in hydrophobic, nonswellable matrices. Examples include stearic acid, lauryl, cetyl or cetostearyl alcohol, glyceryl behenate, carnauba wax, beeswax, candelilla wax, microcrystalline wax and low molecular weight polyethylene, to name a few. Insoluble polymers include fine powders of ammoniomethacrylate copolymers (Eudragit RL100, PO, RS100, PO), polyvinyl acetate or its mixture with povidone (Kollidon SR), ethyl cellulose (Ethocel), cellulose acetate (CA-398-10), cellulose acetate butyrate (CAB-381-20), cellulose acetate propionate (CAP-482-20), and latex dispersions of the insoluble polymers (Eudragit NE-30D, RL-30D, RS-30D, Surelease). However, their applications as the sole rate-controlling material in the marketed matrix-based products are still limited.

19.2.3.2 Materials for reservoir systems

The common materials to form a drug release barrier surrounding a core tablet, drug particles, beads or pellets for diffusion-controlled reservoir systems include water-insoluble acrylic copolymers, ethylcellulose and polyvinylacetate. These film-coating polymers had historically been used in an organic solution prior to the 1980s. Today, they are applied as aqueous dispersions that form films by a process of coalescence of submicron polymer particles in most cases except where the solvent coating is required. Ammoniomethacrylate copolymers (Eudragit RL 30D, RS 30D) are water permeable and swellable film formers based on neutral methacrylic esters with a small proportion of trimethylammonioethyl methacrylate chloride. Methacrylic ester copolymers (Eudragit, NE30D) are neutral esters available as a 30% aqueous dispersion without the need of plasticizers unless increased film flexibility is desired. Ethylcellulose for film coating is available as an aqueous polymeric dispersion containing plasticizers under the brand name of Surelease and as pseudolatex dispersion, Aquacoat ECD, which requires an addition of plasticizers to facilitate film formation during coating. More recently, a polyvinylacetate aqueous dispersion (Kollicoat SR 30D) has been shown to offer similar functionality while providing improved resistance to an aqueous medium containing alcohol in dissolution test.

19.2.3.3 Materials for osmotic pump systems

Cellulose acetate is the most commonly used polymer that constitutes the semipermeable membrane of an osmotic pump device. Cellulose acetate containing a controlled percentage of acetyl content is often used together with other pH-dependent or pH-independent polymers to form films of varying water flux and permeability. Polymers used for such purpose include cellulose butyrate, polyurethane, ethylcellulose, poloxamer polyols, PEG, PVC, and PVA. Osmotic agents, such as sodium chloride, are another key ingredient of an osmotic system. In the more sophisticated two-chamber system, a hydrophilic polymer with a very high swelling ratio; that is, poly(ethylene oxide), is an essential component in the osmotic driving element, the push layer.

19.2.3.4 Materials for delayed release systems

Materials that have been applied to oral medication for enteric release primarily comprise of (1) shellac and zein of natural sources and (2) derivatives of cellulose and methacrylic acid copolymers containing carboxylic functional groups. Technology advances over

the past three decades have made synthetic and semi-synthetic polymers the most preferred system for enteric film coating of tablets or particles. At low pH in the stomach, the polymer is impermeable and insoluble, but in the higher pH environment present in the small intestine (eg, >5.5), the polymer dissolves and exposes the core tablet for the rapid or slow release of the actives. Examples of these pH-dependent anionic polymers include cellulose acetate phthalate (CAP), hydroxypropyl methylcellulose phthalate (HPMCP), methacrylic acid and methacrylic esters (Eudragit L 100, L 100-55, L12.5, L30D-55, S 100, S 12.5, FS 30D, Acryl-EZE), and polyvinyl acetate phthalate (Sureteric). Some of these polymers (eg, Eudragit) are often used to obtain predetermined lag times of the delayed release via pH and/or coating thickness control for colonic and pulsatile (later pulses) release. It should be noted that these ionic polymers have also been incorporated into the matrix, reservoir or osmotic pump systems to modify drug release profiles as discussed in this chapter. For delayed release system based on slow eroding mechanism, materials such as those utilized in matrix systems are used in compression-coated tablets that provide predetermined lag (eg, Adalat CC).

19.3 RATIONAL DESIGN OF MODIFIED RELEASE SYSTEMS

The majority of the systems for modifying apparent drug absorption are, in general, based on proprietary or nonproprietary polymeric delivery technologies. The ability to achieve the desirable in vitro and in vivo performance for a given drug substance is highly dependent upon several important factors that include the dose, physicochemical, biopharmaceutical, pharmacokinetic, and pharmacodynamic properties of the drug, as well as the proper selection of a formulation approach and formulation design. Each drug substance possesses inherent properties that require considerations specific to both the drug and the delivery system. Thus, successful dosage form development is, in fact, dictated by the properties of a compound in the context of physiological/biological constraints, rather than by the technology platform. This conclusion is supported by the fact that almost all MR products with the expired composition of matter patents have been unable to maintain market exclusivity solely based on delivery technology. There are numerous examples where the performance of the branded products has been matched by their generic counterparts based on similar or different ER technologies. Notable examples include Procardia XL, Cardizem CD, Concerta, Adalat CC, Wellbutrin XL, Ditropan XL, Glucotrol XL, Glucophage XR, Asacol, Toprol-XL, to name a few. Therefore, the first stage in designing an MR delivery system should be to conduct technical feasibility assessment by integrating a defined clinical rationale with the characteristics of the drug. The second stage is to select an appropriate MR technology based on the desired dosage form characteristics and other development considerations discussed in chapter "Product and Process Development of Solid Oral Dosage Forms" for in vitro and in vivo evaluation. More specifically, a rational design process should include the following steps:

1. Identify the clinical need and define target in vivo product performance
2. Conduct a study through experiments and detailed analysis to assess feasibility of intestinal absorption, challenges, and risk associated with MR delivery based on
   - Characterization of the molecule concerning its physicochemical and biopharmaceutical properties, dose, regional absorption, in vivo disposition and pharmacodynamics.
   - Pharmacokinetic simulation to calculate theoretical drug input rate and its range required to produce the desired plasma concentration-time profile based on in vivo disposition parameters and certain assumptions.
   - Determine whether physicochemical properties and absorption characteristics of the API in different segments of the GI tract allow for the required drug input within the residence time of the dosage forms (eg, solubility/dose ratio, apparent permeability, single unit versus multiparticulates)
3. If feasible, design and evaluate formulations with different release rates by selecting an appropriate MR technology, associated dosage form, manufacturing process and in vitro test methods.
   - Select prototype formulations with profiles that bracket the theoretical target absorption profile for testing bioavailability in vivo
   - Identify a formulation with acceptable in vivo performance or a direction to aid subsequent round of formulation refinement if iteration is required.
   - Explore an in vitro-in vivo relationship (IVIVR) or correlation (IVIVC) if different prototypes exhibit different in vivo performance.

In this section, the essential design elements and general considerations will be discussed in detail. Case studies are provided to illustrate how drug property, delivery technology and formulation approach influence the performance of oral MR delivery systems.

19.3.1 Identification of the clinical need and definition of the target in vivo product performance

The basic clinical rationale of developing a MR product is to achieve optimal drug treatment via programming the input profile such that the product offers one or more advantages in efficacy, safety, and patient compliance. According to the regulatory directive issued by European Medicines Agency (EMA),[142,143] development of MR products should be based on (1) a well-defined relationship between the pharmacological/toxicological response and the systemic exposure to the drug/metabolite(s) forms and (2) an integration of physiological, pharmacodynamic and pharmacokinetic considerations. In other words, modulating the rate, kinetics, timing and/or site of drug release in the GI tract requires an understanding of the time course of drug concentrations in the body (PK) and its relationship with the pharmacological effect including time course and intensity of therapeutic and adverse effects (PD). However, emphasis has been, in many cases, placed on reducing the dosing frequency or fluctuation of plasma concentrations associated with IR formulations than on understanding and ensuring improved pharmacological effect profile. As a result, it is not uncommon to see new drug delivery systems looking for a suitable drug candidate or projects being terminated due to a lack of feasibility of or the desired outcome. This can be attributed to the often assumed or overly simplified linear PK/PD relationship, and to the fact that the PK outcomes of drug delivery are easy to monitor and quantify. In reality, the relationship between systemic drug concentration (C) and intensity of effect (E) is very complex because of the vast array of pharmacological mechanisms and physiological processes that control drug responses. For instance, the common design objective of an ER dosage form is to maintain the plasma drug concentrations above the minimum effective concentration (MEC) and below the minimum toxic drug concentration (MTC) for drugs with defined therapeutic window. However, for antibiotics with known minimum inhibitory concentration (MIC), there are three types (or indices) of PK-PD relationship that define concentration needed for maximal bacterial killing: (I) concentration-dependent with prolonged persistent effect, (II) time-dependent with minimal persistent effect and (III) time-dependent killing with moderate to prolonged persistent effect.[144–146] The rate of killing is determined by either the length of time necessary to kill or the effect of increasing concentrations. Persistent effects include the Post-Antibiotic Effect (PAE) which is the persistent suppression of bacterial growth following antibiotic exposure. In designing MR dosage forms, the PK target profiles need to differ depending on the PK-PD indices[147] of different classes of antibiotics to maximize clinical efficacy and tolerability. For Type I antibiotics, achieving high AUC_{24h}/MIC and C_{max}/MIC ratios is more important for antibiotic efficacy. For example, a C_{max}/MIC ratio of at least 8–10 is needed to prevent resistance of aminoglycosides. For fluoroquinolones, $AUC_{24\,h}$ /MIC ratios of approximately 125 and 40 are optimal against gram negatives and positives, respectively.[144] For Type II antibiotics, the design of MR products should aim at maximizing the duration of exposure because the cumulative percentage of a 24-hour period that the concentration is above MIC ($T_{>MIC}$) is known to best correlate with efficacy. For beta-lactams and erythromycin, maximum killing is seen when the time above MIC is at least 70% of the dosing interval. Type III antibiotics exhibit time-dependent killing and moderate persistent effects.[144] Thus, the MR product design needs to strive for obtaining high $AUC_{24\,h}$ /MIC ratio with adequate $T_{>MIC}$.[148]

A variety of models based on the mechanisms of drug actions (reversible, irreversible and tolerance) have been developed to describe PK/PD, such as sigmoid E_{max}, biophase distribution, slow receptor-binding, turnover, indirect response, signal transduction and tolerance models.[149] For example, a sigmoidal E_{max} model or the full Hill equation is a four-parameter direct response model:

$$E = (E_{max}\, C^{\gamma})/(EC_{50}^{\gamma} + C^{\gamma}) \qquad (19.21)$$

E is the intensity of the effect, E_{max} is the maximum effect, C is the concentration, EC_{50} is a sensitivity parameter representing C producing 50% of E_{max}, and γ is the Hill or sigmoidicity coefficient that reflects the steepness of the E-C curve. Lower values (<1) will produce broad slopes, higher values (>4) are reflective of an all-or-none type of response. According to this equation, different fluctuations in pharmacological response (ΔE) may result from the same or similar concentration changes (ΔC) depending on the concentration ranges. Additionally, the slow and constant mode of drug input often provided by ER dosage forms may have varying therapeutic impacts depending upon drugs and their PK-PD relationships. Furthermore, the time course of drug effects can be fairly different from the time course of drug concentrations. A short elimination half-life does not necessarily indicate a short duration of action.

Fig. 19.9 shows the sequence of steps between drug administration and clinical response. The importance of efficacy and safety implications of the complex relationship between drug input and pharmacodynamics and research strategies for guiding the design of drug

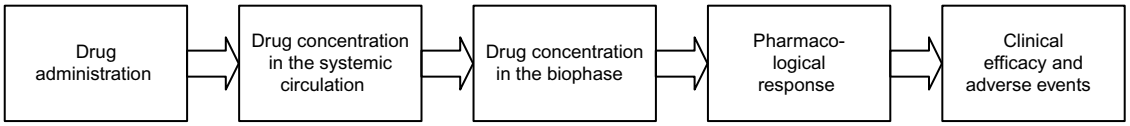


FIGURE 19.9 From drug administration to clinical response.

delivery mode have been discussed in details by Breimer and Hoffman.[150,151] For instance, the shape factor (γ), which determines the slope of the E-C curve of the E_{max} model, can have a significant impact on the design of an ER system. A relatively shallow E-C profile (small γ) may suggest a lack of pharmacodynamic rationale to develop an ER product because E is relatively insensitive to even a large change in concentration. When the concentration range required to elicit drug response is very small (large γ), such as for levodopa, the E-C profile represents an all-or-none phenomenon. Thus, as long as the drug concentrations remain above the MEC, the drug effect is essentially independent of the fluctuation in plasma levels. It is known that the nonsteady state E-C curves of certain drugs may exhibit clockwise or counter clockwise hysteresis due to desensitization or equilibrium delay or formation of an active metabolite.[1] In the former case, either pulsatile or constant rate input may be designed depending upon the desired pharmacological responses of the drugs (eg, gonadotropin-releasing hormone). On the other hand, a slow input rate may minimize the counter clockwise hysteresis and thus require lower C_{max} to achieve the desired response.

The impact of input rate on the efficacy and safety ratio is best illustrated by nifedipine. Comparative studies of slow versus rapid input have led to the realization and confirmation that the rate-of-increase of nifedipine concentration, rather than absolute concentration, is the determining factor for the observed hemodynamic effects; that is, a gradual decrease in blood pressure was observed without side effects (increase in heart rate) occurring with a slow regimen while the opposite was observed with a rapid regimen.[119] Methylphenidate represents another classical example where the delivery pattern can have a significant impact on the efficacy. To meet the clinical need for a more convenient dosing regimen of this controlled drug, several new MR products with different release modes were designed and clinically proven to be effective since the mid-1990s. The product designs include pulsatile release and biphasic extended release that provide fluctuations of plasma levels during drug release, thus overcoming the acutely acquired tolerance due to a constant level of drug exposure associated with Ritalin SR, the extended release version of Ritalin product lines that had been available for many years. These two examples show that a positive or negative clinical outcome may be rendered by constant PK profiles depending on the relationship between the kinetics of drug effects and the pattern and timing of drug input. Therefore, the design of release characteristics for an MR product should be based on the desired optimal drug concentration-time profile; that is, target product profile, defined by the quantitative information of the PK/PD relationship.

With increased research and understanding of biomarkers and surrogate endpoints that reflect the underlying mechanism of drug action, it is anticipated that understanding PK/PD relationships will play a more significant role in the establishment of clinical rationale, a definition of in vivo target product profiles and early decision in MR product development.

19.3.2 Feasibility study

In developing MR dosage forms, feasibility assessment is crucial to product design and development success. Once the delivery mode and in vivo target product profile are defined and disposition parameters are available, the corresponding theoretical drug input profile can be obtained by prospective PK simulation; for example, via adjusting input rate and kinetics to generate in vivo PK profiles, or by deconvolution of the target plasma profile. With a known therapeutic window or MEC, the required input or absorption duration and kinetics can be readily determined, for instance, by ensuring the plasma levels of the MR design constantly above the MEC over the dosing interval. When the quantitative PK-PD information is unavailable, the steady-state C_{min} of the clinically effective IR counterpart may be used as a conservative estimate of the MEC. For example, Fig. 19.10 shows a new ER dosage form was designed to maintain the average plasma concentration at or above the C_{min} (about 100 μg/mL) of a marketed IR dosage form throughout the dosing interval of 24 hours.

Following the completion of the theoretical MR design or paper exercise, the next step would be to do a "reality check"; that is, to evaluate the feasibility of the designed delivery in the GI tract because favorable and complete drug absorption throughout the GI tract is often an exception rather than a rule. For example, if a predefined plasma profile of a drug corresponds to 16 hours of zero-order drug input, it suggests that majority of the administered dose will have to be

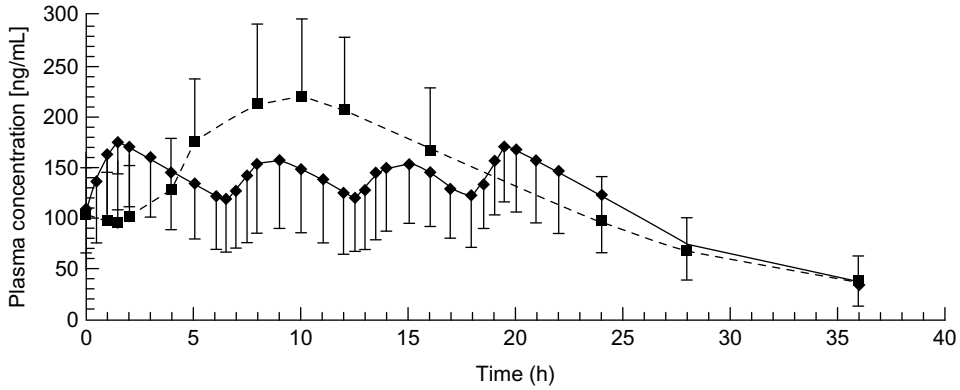


FIGURE 19.10 Mean (±SD) plasma concentration-time profiles of (+)− Tramadol after multiple-dose administration of an immediate release dosage form and a once-daily extended release dosage form in the fasting state (key: 50 mg IR (♦); 200 mg ER (■)). *Reprint with permission from Blackwell Publishing: Malonne, H. et al., Pharmacokinetic evaluation of a new oral sustained release dosage form of tramadol. Br J Clin Pharmacol 2004;57(3):270−278.*

absorbed in the large intestine under fasting condition because of the limited residence time of solid dosage forms in the small bowel. Thus, understanding absorption characteristics of the individual molecule of interest in the lower GI tract is crucial to the rational design of an MR dosage form. To assess technical feasibility and development challenges, it is essential to integrate the drug's physicochemical and biopharmaceutical properties in the context of the total dose, physiological/biological constraints and disposition kinetics. Among the key properties of a particular drug substance are solubility, lipophilicity, chemical stability, regional permeability, gut and hepatic first-pass metabolism, whether or not the drug is a substrate of gut enzymes and transporters and susceptibility of the drug to metabolism by gut microflora. To help better understand the suitability of a particular drug for MR delivery, some of the important factors and their impact on the product development are summarized in Table 19.3.

Orally administered drugs may be subject to acid or base hydrolysis and other chemical or enzymatic degradation that can influence bioavailability and design of the delivery system. For example, drugs that are unstable in the stomach require delayed onset of drug release until the dosage form enters the small intestine. This may result in shortened total GI residence time for extended absorption. On the other hand, drugs unstable in the lower GI tract due to metabolism by microflora would require complete drug delivery in the small intestine.

The regional absorption characteristics of a compound in the GI tract and residence time of dosage forms are the most important parameters in assessing the suitability of oral MR delivery.[1,153] With significant regional differences including surface area, permeability, secretion, enzymes, transporters, the volume of water, and so on, a drug substance often exhibits varying degrees of apparent absorption in different segments of the GI tract. In general, the transit time of most dosage forms before arrival at the colon, a less favorable site for drug absorption, is approximately 4−6 hours depending on the type of dosage forms, food intake, and composition. This may become a limiting factor for drugs requiring absorption beyond this time frame after dosing.[154] For certain drugs, favorable absorption in the large intestine can allow continued drug delivery for up to a total of 20−24 hours.[155,156] If drug release is not completed by the time the dosage form passes through the absorption region, a portion of the dose will not be delivered. Hence, it is very important to define the absorption regions or window of a specific compound in the GI tract before further development proceeds. MR dosage forms designed and tested in adults may not work the same with the young children because of the known differences in GI transit time of the delivery systems between adults and children younger than 3−4 years old. With shorter GI residence time in young children, truncation of drug absorption may occur depending on the characteristics of the drug substance and MR design, resulting in lower AUC and/or subtherapeutic plasma levels.[157]

Over the years, many techniques have been utilized to assess regional absorption potential or window of a compound, including in vitro or in-situ models and site-specific delivery in animal models and human subjects. Permeation characteristics of drugs in different GI segments can be evaluated through in vitro permeability of excised tissues or in situ perfusion studies in animal models (eg, rat or rabbit). Site-specific delivery via indwelling access ports in conscious animals (eg, dog and rat) offers a direct comparison of drug absorption in jejunum, ileum, and colon through

TABLE 19.3 Factors Affecting the Feasibility of Developing Oral Modified Release Systems

A. SOLUBILITY/DOSE
Low solubility combined with high dose may limit the suitability for developing an MR system that requires drug delivery in the lower bowel. Drug absorption is usually more variable or incomplete when the in vivo release occurs past the ileocecal junction where the amount of fluid available for dissolution is progressively limited, and the permeability and surface area are lower. In some cases, apparent absorption of poorly soluble drugs may be inherently prolonged due to the slow dissolution of the drug particles. However, achieving consistent drug release requires tight control of bulk powder properties of the API, including particle size distribution, morphology, surface properties and agglomeration, etc.
B. STABILITY
Drugs must be stable to pH, enzymes and flora throughout the entire intended delivery regions of the GI tract. For example, a drug subject to degradation by the colonic microflora is not a suitable candidate for MR delivery that requires an in vivo absorption beyond approximately 5–6 hours post dosing.
C. LIPOPHILICITY/PERMEABILITY
Absorption of BCS Class 3 or 4 drugs may be limited by permeability that also differs with regions of the GI tract due to difference in available surface area, uptake and/or efflux transporters involved.[152] Altering the release of such drug substances in the lower GI tract typically have little effects on the shape of PK profiles and usually result in truncated absorption when > 5–6 hours of in vivo absorption is required.
D. ELIMINATION $t_{½}$
The need for ER systems is often a result of short half-life. However, other variables (eg, MEC, a volume of distribution and dose) are also important in determining the ER feasibility. Hence, developing a twice-daily (or once-daily) system may be possible for one compound, but not feasible for another drug having similar half-lives.
E. THERAPEUTIC WINDOW
One of the common design objectives of oral ER drug delivery is to maintain plasma levels within the therapeutic window with reduced fluctuation. For drugs with relatively short half-lives, the lower the MEC and the higher the minimum toxic concentration (MTC), the more likely it is to achieve prolonged drug exposure above MEC using higher dose. However, this will result in a greater fluctuation in steady-state plasma levels that is often undesirable for narrow therapeutic index drugs.
F. FIRST-PASS METABOLISM
For drugs with saturable first-pass metabolism (hepatic or gut), bioavailability may be compromised due to slow input if drug absorption from IR dosage form is higher because of transient saturation. No significant impact is expected if drug absorption of both IR and ER is within linear range (ie, dose proportional).
G. PK-PD RELATIONSHIP
The relationship between drug concentration (C) and the pharmacological effect (E) plays a critical role in determining rationale and need for MR delivery as previously discussed.

concurrent monitoring of plasma levels. It can serve as a predictive model for human exposure depending on the drug's properties and absorption mechanism. In fact, a reasonably good correlation of regional absorption characteristics has been found between dog and human, particularly for passively transported compounds.[158] For small hydrophilic molecules for which the paracellular pathway also plays a role in permeation, the dog model usually overestimates the absorption in humans due to the "leakier" canine intestine, especially in the large intestine. To quantify colonic absorption for delivery potential, intubation studies have also been used via bolus and/or continuous infusion of a drug solution or suspension into the transverse colon in animal or healthy volunteers. Since the early 1990s, gamma scintigraphic studies have become one of the commonly used techniques for screening drug candidates for oral MR delivery.[159] For instance, regional differences in absorption can be determined by using noninvasive delivery devices, such as the Enterion capsule of Quotient Clinical[160] and the InteliSite capsule of Scintipharma, Inc. These radiolabeled capsules loaded with a drug solution or powder can be tracked via gamma scintigraphy and externally activated to release the drug when it reaches the desired location in the digestive tract following ingestion. Another effective method used by the two companies and Bio-Images Research Ltd. is to administer a radiolabelled MR dosage form followed by combined analysis of its GI transit characteristics along with location and timing of drug release.[161] Recently, Medimetrics developed a new system known as IntelliCap that combines the functions of drug delivery, measurement, and wireless communication.[162,163] It is an electronic drug delivery capsule that enables programmable dispensing of a drug while continually measuring and reporting pH and temperature as it travels through different segments of the GI tract. Based on the measured pH profile the capsule location is determined and individual GI transit times are obtained from each capsule. Hence, concurrent scintigraphic imaging is not required. Two product versions are available for controlled delivery of fluids, solids or powders: (1) IntelliCap CR is constructed for a fluid payload and dispenses drug in small increments under control of the microprocessor, allowing the creation of

a variety of delivery profiles; (2) IntelliCap FR is designed to deliver drug as a single fast-release event for flexible payload type including powder, particles, solid fluid. All of these studies are useful in evaluating absorption characteristics of the drug molecules in different segments of the intestinal tract with minimum confounding factor. However, the oral delivery feasibility of MR systems not only depends on regional absorption properties of the API but also variables associated with dosage form design, composition, and/or processing. Hence, an alternate approach that can often be efficient and effective is to incorporate feasibility assessment as part of the early product development. By applying interdisciplinary knowledge and integrating the drug's properties (physicochemical, biopharmaceutical and biological) with regional absorption information in animal models, a decision can often be made initially with regard to proceeding with formulation screening study in humans. The first screening study typically involves administering prototype formulations with varying in vitro release profiles that bracket the simulated target in vivo absorption profile by first assuming a wide window of absorption. If the study indicates decreasing bioavailability with slowing drug release, the absorption window can be inferred or approximated by deconvolution or examination of plasma profiles (eg, using a semilog plot). For example, the limit of drug release duration in vivo normally corresponds to where the observed truncation of apparent absorption occurred. The reliability of the information thus obtained usually depends on the number of the prototypes tested, the IVIVR, and understanding of certain potential confounding factors. This approach is appealing because (1) the time and cost required for a screening study are similar to that of a "cleaner" scintegraphic study and (2) with sufficient rationale, knowledge and experience invested in the design of prototypes and in vitro screening tools, there is a reasonable chance of identifying a target formulation in the same study or obtaining a rank order of in vivo performance which can often be used to guide formulation modification when necessary. The main drawback of this approach is its potential impact on making a "go/no go" decision when the study result is inconclusive with respect to the absorption window due to confounding factors.

The therapeutic window and MEC and their interplays with disposition characteristics of a drug also play a critical role in suitability evaluation for ER delivery. In theory, a lower MEC combined with slower elimination corresponds to a shorter required duration of drug input. Drugs with very short half-lives ($t_{1/2}$) and high volume of distribution undergo rapid clearance, which can make extended delivery challenging especially if the MEC is relatively high or therapeutic window is narrow. For instance, prolonged absorption is often required to achieve essentially flat plasma profiles for drugs with short half-lives and narrow therapeutic indices since fluctuation of the plasma profiles depends on the rates of elimination and input duration/kinetics. Therefore, a wide absorption window is often necessary to achieve extended release of this type of drugs. Lastly, it should be noted that a feasibility study is not required for developing generic MR products because it is already confirmed by the presence of branded products.

19.3.3 Selecting the MR system and testing system design

When oral MR delivery is feasible or shows promise, the logical subsequent step is to choose an appropriate MR technology that is capable of meeting a specific delivery need along with in vitro tests for evaluating prototype formulations with different release rates in vitro. This should be built on the understanding of drug properties, required dose and release/absorption kinetics, clinical and market needs including type and size of dosage form, number of strengths, among others. Additional practical considerations include development time and cost, and commercial production factors discussed in chapter "Product and Process Development of Solid Oral Dosage Forms" (eg, process, equipment, facility, manufacturability, robustness, cost, capacity, environment, and so on). In selecting an MR system that matches development needs, dose and solubility of the active are often the most important factors to consider because, within the size constraint of oral dosage forms, both variables impact the release mechanism and processing behaviors. Thus, there is no technology that is "one-size-fits-all." Nevertheless, it is not uncommon for a drug delivery company to tout a specific technology while downplaying the practicality and viability of technology transformation, such as system and process complexity, and many other factors discussed previously. Table 19.4 compares delivery capabilities and limitations of commonly utilized ER systems. From an enablement point of view, different MR systems can be equally effective in achieving the delivery objective when drug properties and dose are desirable. For example, similar MR performance of verapamil, theophylline and methylphenidate have been obtained using matrix, reservoir or osmotic pump technologies. Depending on dose and solubility, one type of delivery system may become more or less suitable for meeting a particular delivery need. Table 19.5 provides a high-level guide to the selection of ER systems on the basis of dose and solubility.

TABLE 19.4 Comparison of Commonly Used Oral Extended Release Technologies

System	Advantages	Disadvantages
Hydrophilic matrix	Suitable for compounds with a wide range of properties and low to high drug loading Generally robust formulation and process when rationally designed Use of conventional manufacturing equipment and process Cost effective: Shorter development time and lower cost Release kinetics and profiles can be tailored with modification Multiunits possible	Drug release often sensitive to test conditions Less flexibility to adjust dose strengths for single-unit system Increased formulation/process complexity to achieve tailored drug release (layered or compression coated system)
Hydrophobic matrix	Suitable for soluble compounds and low to high drug loading Use of conventional manufacturing equipment and process Release kinetics and profiles can be tailored with modification Multiunits possible	Not applicable to compounds with low solubility Nonzero-order release Propensity for incomplete drug release Drug release often sensitive to processing and test conditions Less flexibility to adjust dose strengths for single-unit system Increased formulation/process complexity for tailored drug release (layered or compression coated system)
Multiunit reservoir	Readily tailored release kinetics and profiles (eg, zero-order, pulsatile, biphasic, colonic) Minimized risk of dose dumping and local irritation Lower in vivo variability owing to favorable GI transit properties More consistent in vivo performance Easy dose adjustment: Single formulation amenable to multiple strengths Suitable for pediatric/geriatric use Use of conventional manufacturing equipment and process	Only applicable to compounds with relatively high solubility Drug release often sensitive to test conditions Limited drug loading Process development and scale-up more challenging
Osmotic pump	Applicable to compounds with a relatively wide range of properties Drug release generally independent of drug properties and test conditions Zero-order release	Limited drug loading Ghost tablets Delayed onset (1-2 h) and/or incomplete drug release Solvent-based process Lengthy, complex and inefficient manufacturing processes and control (eg, Layered tablet) Specialized equipment, facility Highest development and manufacturing cost and time

TABLE 19.5 Guide For ER System Selection on the Basis of Dose and Solubility

System	HS/HD	HS/MD	HS/LD	MS/HD	MS/MD	MS/LD	LS/HD	LS/MD	LS/LD
Hydrophilic matrix tablet	0	+	+	+	+	+	+	+	+
Hydrophobic matrix tablet	+	+	+	–	0	+	–	–	–
Hydrophobic matrix pellets	–	–	–	–	+	+	–	+	+
Coated matrix tablet	+	+	+	–	0	+	–	–	–
Coated pellets	–	+	+	–	+	+	–	–	–
Osmotic pump	–	0	+	–	+	+	–	+	+

HS, high solubility; MS, medium solubility; LS, low solubility; HD, high dose; MD, medium dose; LD, low dose; + , suitable; − , unsuitable; 0, borderline (may be suitable via system modification).

Once an MR system is chosen, in vitro characterization of its behavior and attributes is essential prior to testing in human. Factors that often affect the system performance such as, type and size of the dosage form, drug loading, excipient property, processing method, release mechanism, sensitivity to environmental changes, need to be considered when choosing appropriate in vitro test methods. For solid MR dosage forms, drug release test is the most important among various physical and chemical characterization methods. The USP drug release test (eg, using apparatus II or I) is commonly utilized for gauging the likely in vivo performance, thus guiding formulation screening despite its undefined relationship with in vivo performance at the early design stage. Depending on the drug properties and system design, test parameters including apparatus, agitation, pH, surfactants, additives, ionic strength, mechanical stress, among others, can be varied to better mimic certain specific GI conditions or to create shear stress on the structure for testing the performance robustness. A detailed discussion of development considerations related to meaningful in vitro release tests is provided in chapter "In Vitro/In Vivo Correlations: Fundamentals, Development Considerations, and Applications."

In designing and evaluating an MR dosage form, the initial objective is usually to prepare a formulation with an in vitro release profile similar to the theoretical in vivo absorption profile. Due to unknown IVIVR, additional formulations having in vitro profiles that bracket the target profile are useful for identifying a formulation with desired in vivo performance when they are tested in humans for the first time. In the event where an iterative study is required, the data from such a screening study provide not only a direction to formulation refinement but also an opportunity for early exploration of IVIVR or IVIVC. As discussed in chapter "In Vitro/In Vivo Correlations: Fundamentals, Development Considerations, and Applications," the earlier an IVIVC or IVIVR can be established, the more beneficial it is to the design and development of a MR product. A validated IVIVC offers a useful scientific and regulatory tool to aid formulation and process development and scale-up, justify specifications and make changes related to process, equipment and manufacturing site. Lastly, with any technology chosen, it is imperative to integrate the selection of formulation components and associated manufacturing process based on the multifaceted fundamental understandings discussed in chapter "Product and Process Development of Solid Oral Dosage Forms" in order to ensure consistent quality and a reproducible performance in vivo. For example, oral dosage forms are subjected to a range of shear forces such as peristalsis and also encounter a variety of pH and chemical environments during GI transit. In poorly formulated hydrophilic matrices, these mechanical and chemical challenges can potentially cause the matrix to lose its integrity prematurely and break up.[30] Thus, product design requires consideration of material properties, process operation and possible interactions that may influence the matrix integrity and consistent drug release. In particular, knowledge of the basic characteristics of the rate-controlling polymer and fundamental factors that affect drug release[30] is essential to design hydrophilic matrices with robust performance.

19.3.4 Case studies: impact of drug property and formulation design

The examples discussed in this section are intended to illustrate (1) the criticality of understanding drug characteristics and feasibility assessment and (2) there is no one-size-fits-all solution to designing an MR product (3) delivery technology is inconsequential for drug molecules with desirable properties for MR development.

19.3.4.1 Case study 1: methylphenidate HCl

Methylphenidate (MPH) is an amphetamine-like central nervous system stimulant commonly prescribed to treat ADHD in children, adolescents, and adults, as well as narcolepsy. It is a weak base with pK_a of 8.77 and a log$P$ of 3.19. Its hydrochloride salt is freely soluble in water (18.6 mg/mL), stable and well absorbed from the intestinal tract with a short elimination half-life of 3–4 hours.[164] These favorable properties combined with a low dose make MPH an ideal candidate for oral MR regardless of technologies applied.

The IR product (eg, Ritalin) that has been on the market since the 1950s is taken two or three times daily with rapid onset of action and demonstrated efficacy. In an attempt to offer more convenient administration and prolonged control of behavior, especially for school children for whom most prescriptions are written, several modified-release formulations have been developed with the objective of maintaining efficacy over the entire school day. This is of particular importance because it would eliminate the need for repeated administrations of methylphenidate, a Schedule II controlled substance while the patient is at school. An early MR product, for example, Ritalin SR, is a matrix tablet designed to maintain relatively flat plasma drug levels for the required treatment duration. However, the product did not provide equivalent efficacy to the multiple dosing of the IR product during the day due to the tolerance to the effects of the

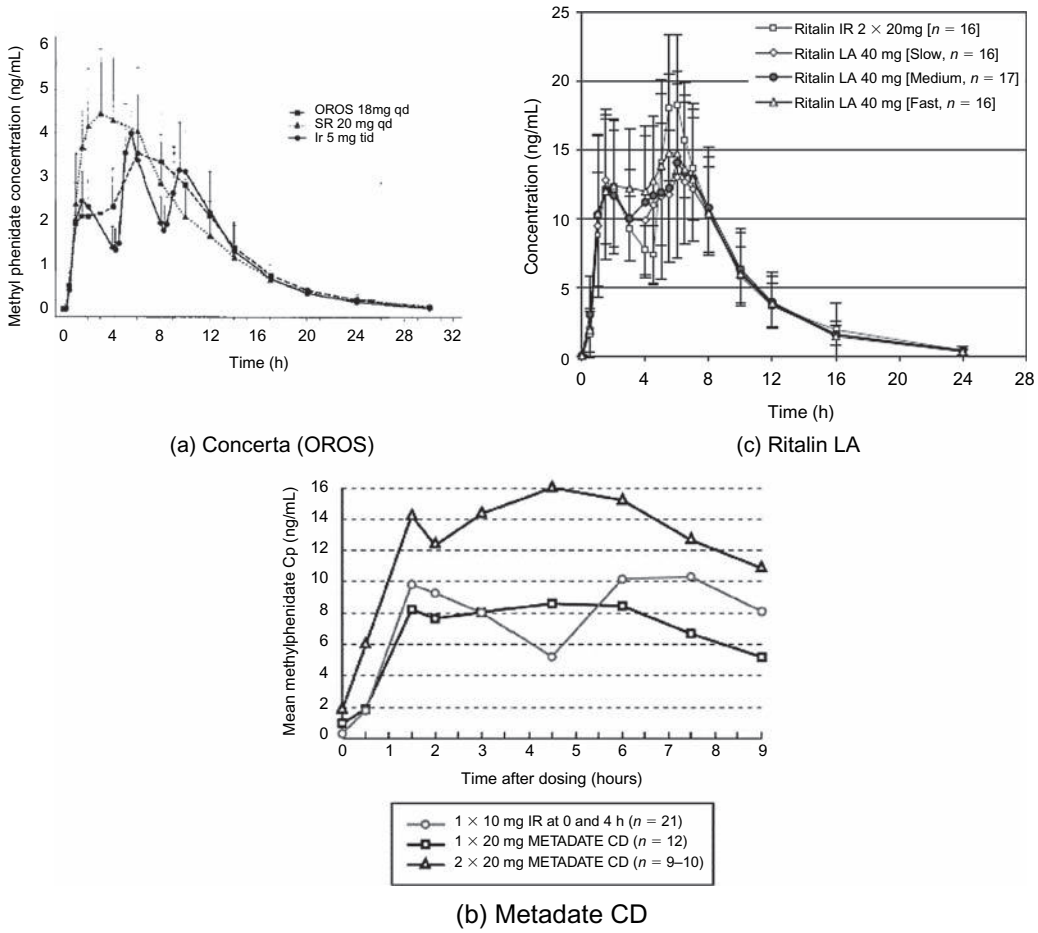


FIGURE 19.11 Plasma Concentration-time Profile of MPH after a Single Dose of MR and IR Formulations (All versus Ritalin SR tablet and/or IR tablet given 4 h apart). *Reprinted with permission: J Clin Pharmacol. 2000;40:379–388; Metadate® CD Package Insert, UCB, Inc.; Biopharm Drug Dispos 2004;25(2):91–8.*

drug discussed previously and a delayed onset of action.[165] Other similar ER matrix tablets include Metadate ER and Methylin ER. It was not until the 1990s when it was demonstrated that varying rate of drug input would be able to overcome the tolerance observed with the formulations that release the drug at nearly constant delivery rate. Subsequently, a new generation of once-daily MR products was developed that offer the equivalent efficacy of repeated administrations of IR MPH by altering the drug's input profile. The successful commercial products include (1) ER products generating biphasic release profiles for a rapid onset of action and a prolonged duration of action, such as Metadate CD capsules consisting of 30% IR and 70% coated ER beads and Concerta tablets consisting of an ER osmotic pump coated with IR component (Fig. 19.11a and b).[166,167] (2) Pulsatile release products that mimic PK performance of standard schedules of IR product, such as Ritalin LA capsules consisting of 50% IR and 50% DR beads that provide IR release 4 hours later (Fig. 19.11c).[168,169] This case study shows that all three types of ER technologies are equally effective in achieving a range of in vivo delivery performance when the molecule exhibits desirable characteristics suitable for MR design.

19.3.4.2 *Case study 2: clarithromycin*

The above example shows that favorable absorption characteristics combined with the rapid elimination of MPH make it possible to achieve the desired PK fluctuation using short pulsing intervals of drug input. To further evaluate the influence of drug's PK properties on the feasibility of oral pulsatile delivery, clarithromycin was selected based on its physicochemical and biopharmaceutics properties that are suitable for MR drug delivery.[170] It is a weak base with pK_a of 8.76 and solubility of 1.2 mg/mL at pH 6.8. It has a log$P$ of 2.69 and adequate stability in the intestinal tract. Its disposition parameters available from the literature indicate a terminal elimination rate constant of 0.115 hour^{-1}. PK simulation of the in vivo performance of the pulsatile release designs with varying pulsing intervals was performed by shifting the

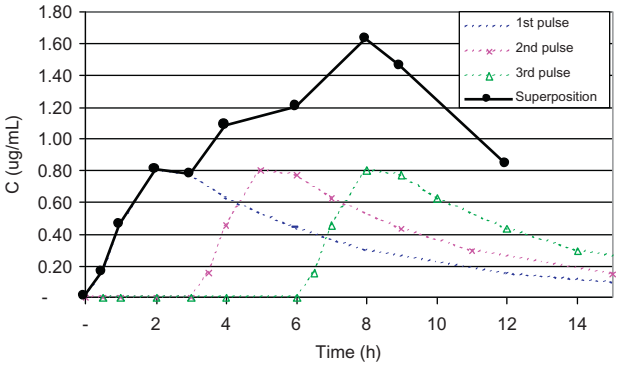


FIGURE 19.12 PK simulation of a pulsatile delivery design by superposition.

observed plasma profiles of a single dose study in humans using superposition principle. The basic assumptions for the simulation included (1) the formulation technology can deliver the exact pulses at predetermined time intervals and (2) the apparent absorption is uniform over the delivery duration with minimum impact of variable GI transit time, gut metabolism and release/absorption environment. An example illustrated in Fig. 19.12 shows that a design of three equally spaced pulses of equal doses over 6 hours hardly generated distinct peaks and troughs in plasma levels corresponding to the input variations. The fluctuations in plasma levels at steady state (not shown) were essentially nonexistent. Thus, the chance is very low for achieving multiple peaks/troughs at the biophase or site of drug action, which is the design objective.[171] It should be noted that this type of simulation represents the best-case scenario because the PK profile used for generating the profiles of the subsequent pulses resulted from absorption that occurs in the most favorable absorption environment in the upper intestinal tract. In reality, the second and third pulsing doses are expected to produce slower ascending plasma concentrations compared to the first pulse since they are absorbed in the distal ileum and/or large intestines. This case study illustrates that applying an in vitro MR design to obtaining a desired in vivo PK and clinical results requires the application of multidisciplinary knowledge and prospective evaluation of in vitro and in vivo outcomes before execution of the experiments. For pulsatile delivery, rapid clearance is essential to achieving distinct PK fluctuation resulting from pulses of shorter intervals. In addition to that, the influence of absorption window and inherently highly variable GI systems on performance of oral dosage forms should also be taken into consideration.

19.3.4.3 Case study 3: development compound A

A development compound in Phase III clinical trials is a neutral molecule with molecular weight of 236 and favorable log$P$ of 1.44.[2] It is stable and practically insoluble (170 μg/mL). The IR formulation requires a q.i.d. dosing regimen due to rapid elimination in humans ($t_{1/2} = 2.5$ hours). Based on the MEC of 1 μg/mL and a total daily dose of 2.4 g, PK simulation indicated that twice-daily administration would require an approximately 10 hours of in vivo input, suggesting at least 50% of the dose would be released/absorbed in the large intestine for a zero-order delivery system. However, regional absorption data were not available. To investigate technical feasibility of ER delivery, prototype hydrophilic matrix tablets and multiunit spherical hydrophobic matrix beads were designed as adult and pediatric dosage forms, respectively. Matrix systems were used because of the required high loading of an API with low solubility. In vitro studies showed apparent zero-order drug release from the matrix tablet due to the erosion-controlled mechanism. Drug release from the matrix beads more closely followed diffusion-controlled kinetics as a result of a combination of high surface area-to-volume ratio, short diffusion path-length and low driving force for drug release. Crossover PK studies of these formulations with varying release rates were conducted in healthy volunteers. The dosage forms were administered under a high-fat condition to increase gastric retention time. Evaluation of the plasma concentration-time profiles showed (1) decreasing release rates correspond to lower AUCs, (2) truncated absorption as indicated by the evident slope inflection of the semilog plot of the terminal phase occurred at approximately 5–6 hours post dosing. Interestingly, the average truncation time was delayed by about 1–2 hours for the matrix beads, suggesting a more favorable GI transit properties of multiparticulate dosage forms. Nevertheless, neither dosage form met the predetermined criteria of maintaining plasma levels above the MEC for approximately 12 hours. Concurrent evaluation of the metabolites showed that the absorption truncation was a result of degradation of the active compound by microflora present in the large intestine, which is consistent with the transit time of dosage forms in the upper GI tract. This case study shows that in vivo delivery window in the GI tract is determined by the inherent property of the drug molecule. Delivery technology is ineffective in overcoming feasibility obstacles when there is a large gap between the technology capability and required oral MR delivery need.

19.3.4.4 Case study 4: development compound B

A development compound in Phase II clinical trials is a polar molecule with molecular weight of 284, a pK_a of 8.3 and a log$D$ of −2.0 at pH 6.8. It is stable and highly soluble (> 0.8 g/mL) in the physiological pH range. Low permeability (Caco-2 $P_{app} = 0.27$ ×

10^{-6} cm/second) suggests it is likely a BCS class III compound. Phases I–II clinical studies used an IR formulation administered every 6 hours due to rapid clearance in humans ($t_{1/2} = 4$ hours). Based on the projected MEC of 10 μg/mL and a total dose of 6 mg, PK simulation indicated that twice-daily administration would require approximately 10 hours of in vivo input, suggesting a significant portion of the dose would be released and absorbed in the large intestine. A preliminary study for assessing the regional absorption in dogs via indwelling access ports to different regions of the intestinal tract showed that AUC values of ileum and colon delivery relative to that of jejunum were 0.94 and 0.67, respectively. In vitro drug release that follows square-root-of-time and near zero-order kinetics were obtained from matrix and ethylcellulose coated multiunit minitablets, respectively. However, a crossover single dose study of three ER formulations with decreasing release rate in healthy volunteers resulted in relative bioavailability values of 50%, 34%, and 29%, respectively. Truncated absorption was evident at approximately 4 hours post dosing under a fasting condition for all three ER formulations, indicating a lack of absorption in the lower bowel. The result of the dog regional absorption model was considered an outlier because an extensive database generated had demonstrated the predictive nature of the model. Combined analysis of drug properties, absorption mechanism and the physiological differences between human and the dog model suggests that the favorable absorption characteristics in the dog colon are likely a result of the leakier dog intestine that favors paracellular transport mechanism of a low dose, small polar molecule. This is consistent with the discussion of Section 1.3.2 and the results of a study that used the dog model for predicting colon permeability in humans. This study found that the relative bioavailability from administration to the colon of the dog correlated well with that of humans for model compounds of all four BCS classes except for atenolol, a hydrophilic small molecule. This case study illustrates the importance of understanding both the property and transport pathways of a compound in assessing in vivo absorption window for oral MR delivery.

19.3.4.5 Case study 5: oxybutynin HCl

Oxybutynin is an antispasmodic, anticholinergic agent indicated for the treatment of overactive bladder with symptoms of urge urinary incontinence, urgency, and frequency by relaxing bladder smooth muscle. It is a weak base with a pK_a of 9.87 and log$D$ of 1.83 at pH 7.3. Its hydrochloride salt is freely soluble in water (12 mg/mL), stable and well absorbed from the intestinal tract. These favorable properties combined with the low dose (5–15 mg/day) make it an ideal candidate for oral MR independent of technologies applied.

For over 20 years, use of IR dosage forms of oxybutynin has been limited by anticholinergic adverse events, such as dry mouth, difficulty in micturition, constipation, blurred vision, drowsiness, and dizziness. These side effects are dose-related and sometimes severe. Studies have indicated that oxybutynin metabolism mediated by the cytochrome P450 3A4 isozyme occurs primarily in the proximal GI tract and liver. Its primary active metabolite, *N*-desethyloxybutynin, is also responsible for much of the adverse effects. After administration of the IR formulation, *N*-desethyloxybutynin plasma levels may reach as much as six times that of the parent drug, substantially affecting the tolerability of the compound within the individual. For instance, in one population studied, after 6 months more than half of the patients had stopped taking the medication due to side effects. Another study estimate is that over 25% of patients who begin oxybutynin treatment may have to stop because of dry mouth.

Since the 1990s, ER products have been investigated and developed (eg, Ditropan XL, Cystrin CR tablets) to allow once-daily administration and prolonged control of symptoms. Clinical studies have shown that ER formulations not only consistently enhanced patient compliance and efficacy but also resulted in improved tolerability including significantly higher continuation rate due to marked reduction of moderate/severe dry mouth. The latter was attributed to substantially lower intestinal metabolism because a majority of the parent drug of the ER dosage form is delivered to the distal intestinal tract where metabolic enzyme activity in the gut wall is lower or absent. As a result, a substantial increase in bioavailability relative to the IR formulation was also imparted. More recently, to totally bypass presystemic metabolism, a 4-day transdermal patch (Oxytrol) was later successfully developed with tolerability profile similar to that of a placebo.[172–174] This example shows that understanding of a drug's disposition properties, biological and pharmacodynamic characteristics enable rational development of MR delivery strategy and design with maximized medical benefits.

19.3.4.6 Case study 6: phenylpropanolamine HCl in EOP

Phenylpropanolamine (PAA) is a sympathomimetic amine that shares structural similarities with amphetamine and ephedrine. The effects of PPA are largely the result of α-adrenergic agonist activity. Before its removal from the market after being around for almost 50 years, it was mainly sold in over-the-counter (OTC) preparations as decongestants or as appetite suppressants. Because of its association with increased risk of hemorrhagic stroke, FDA removed PPA from

OTC sales in 2005. Despite its current ban on human drug use, it is worthwhile to examine it from a historical perspective to illustrate the application of rational MR dosage form design principles to the development of the first osmotic oral dosage form containing PPA as an appetite suppressant in the early 1980s (Acutrim marketed by Ciba-Geigy). PPA is a weak base with a pK_a of 9.4 and a log$P$ of 0.67. Its hydrochloride salt is freely soluble in water and well-absorbed in the GI tract with over 60% bioavailability. It also exhibits linear pharmacokinetics in the dose range 25−100 mg with a plasma half-life of 2−3 hours. Because of its short half-life, an attempt was made to design a membrane-controlled drug delivery system to maintain a constant rate of drug delivery for once-a-day dosing rather than three times a day IR dosing to reduce the fluctuation of plasma levels. As reported by Good and Lee,[175] a desired steady-state plasma PPA concentration of 80 ng/mL was initially selected as the in vivo target (from the known minimum therapeutic plasma concentration range of 60−100 ng/mL) with the requirement that this steady-state level is attained as soon as possible after ingestion. Based on a one-compartment open model using the known pharmacokinetic parameters for PPA ($k_a = 1.58$, $k_e = 0.159$ & $V_D = 245.6$ L from oral solutions), a steady-state drug flux of 3.12 mg/hour from the delivery system was determined to be required to achieve this target steady-state plasma level, and an immediately available dose of 21.6 mg to be needed to quickly establish the desired steady-state level. An EOP for phenylpropanolamine-HCl was then designed to achieve these target levels taking into consideration of additional constraints that the total dose size is limited to 75 mg, and at least 90% of the total dose must be delivered within 24 hours. The final dosage form preserved most of the predicted properties except with some fine tuning, in which 20 mg of the drug is provided as an IR portion as a bolus coating on the EOP, the reservoir of 55 mg to be delivered through the orifice at a constant nominal rate of 3.5 mg/hour for 12.5 hours, and the remainder having a decreasing delivery rate as the excess solid drug in the core is being depleted. The in vitro dissolution profile for this EOP dosage form is compared with the predicted release profile in Fig. 19.13.

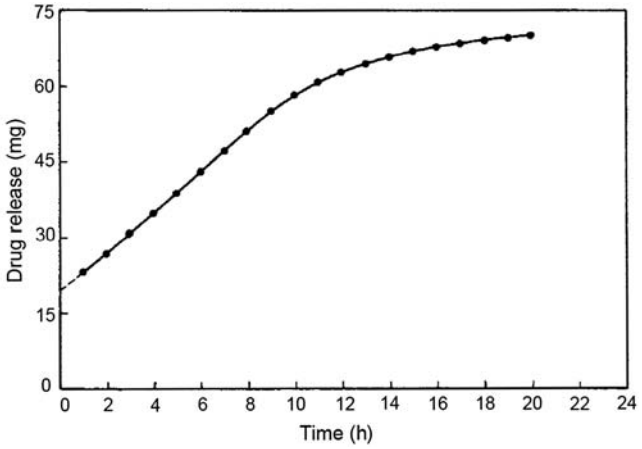


FIGURE 19.13 In vitro dissolution of PPA form EOP dosage containing a 20-mg bolus and 55-mg core. ●, experimental; −, predicted. *Figure adapted from W.R. Good and P.I. Lee. Membrane-controlled reservoir drug delivery systems. In: Langer RS, Wise DL, editors.* Medical applications of controlled release, *vol. I. Boca Raton, FL: CRC Press, Inc.; 1984. p. 1−39 (reproduced with permission from CRC Press, Inc. Copyright, 1984).*

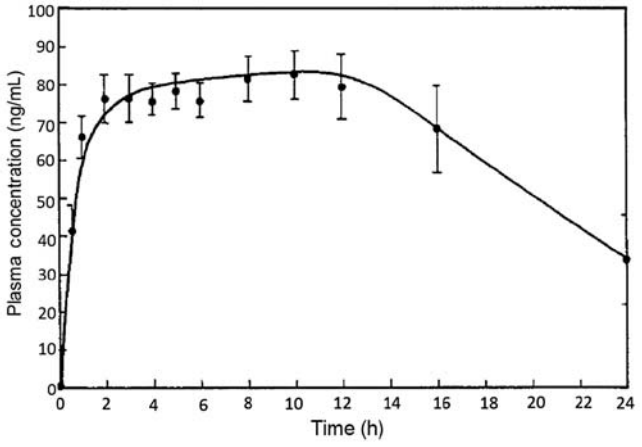


FIGURE 19.14 Comparison of plasma concentration profiles of unchanged PPA after a single dosing of an EOS dosage form containing a 20-mg bolus and 55-mg core ($n = 6$). ●, experimental; −, predicted. *Figure adapted from W.R. Good and P.I. Lee. Membrane-controlled reservoir drug delivery systems. In: Langer RS, Wise DL, editors.* Medical applications of controlled release, *vol. I. Boca Raton, FL: CRC Press, Inc.; 1984. p. 1−39 (reproduced with permission from CRC Press, Inc. Copyright, 1984).*

This EOP dosage form was subsequently tested in a pilot clinical study in six healthy subjects.[175] The resulting mean plasma concentration profile is shown in Fig. 19.14 together with that predicted from the known pharmacokinetic parameters. Additionally, the in vivo release profile deconvoluted from plasma data is compared with the in vitro dissolution profile in Fig. 19.15. It is clear from Figs. 19.14 and 19.15 that the agreement between the experimental data and that predicted from the pharmacokinetic analysis is excellent, particularly the indication of IVIVC in Fig. 19.15 further supports the usefulness of applying the rational design principles described here for a priori MR dosage form design.

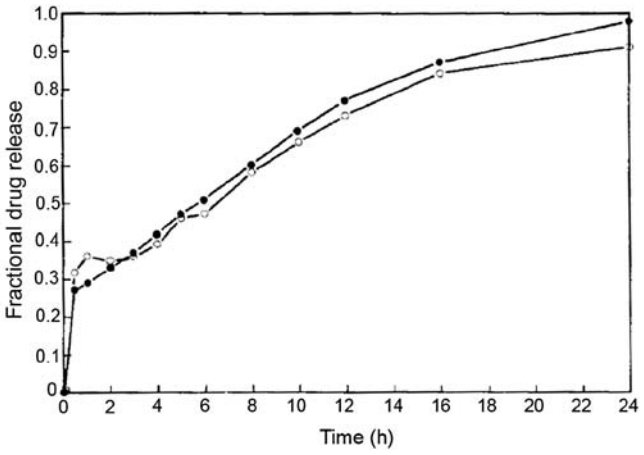


FIGURE 19.15 Comparison of in vitro release profile (○) deconvoluted from plasma data and in vitro dissolution profile (●) for an EOS dosage form containing a 20-mg bolus and 55-mg core. *Figure adapted from W.R. Good and P.I. Lee. Membrane-controlled reservoir drug delivery systems. In: Langer RS, Wise DL, editors.* Medical applications of controlled release*, vol. I. Boca Raton, FL: CRC Press, Inc.; 1984. p. 1–39 (reproduced with permission from CRC Press, Inc. Copyright, 1984).*

19.4 SUMMARY

The most important elements of utilizing MR drug delivery design to achieve desired clinical benefits include defining the clinical rationale and understanding the drug molecule. The success of designing a commercially viable MR dosage form depends on the physicochemical, biopharmaceutical, pharmacokinetic and biological properties of the drug candidate, rather than on a particular delivery technology. Today, almost all oral MR delivery needs can be addressed using at least one of the three broad types of well-established delivery technologies discussed in this chapter. However, there is no "one-size-fits-all" technology. It is more important to select a technology based on technical, practical, developmental, operational and commercial considerations because for many compounds with desirable properties, different "sizes" of technologies will fit "one" molecule.

References

1. Qiu Y, Zhang G. Research and development aspects of oral controlled release systems. In: Wise DL, Klibanov AM, Langer R, Mikos AG, Peppas NA, Trantolo DJ, Wnek GE, Yaszeski MJ, editors. *Handbook of pharmaceutical controlled release technology*. New York, NY: Marcel Dekker, Inc; 2000. p. 465–503.
2. Qiu Y. Design and evaluation of oral modified-release dosage forms based on drug property and delivery technology. Proc. CAE/AAPS/CPA/CRS/FIP/APSTJ Conference: Oral Controlled Release Development and Technology. Shanghai; 2007. p. 34–43.
3. Getsios D, Caro JJ, Ishak KJ, El-Hadi W, Payne K, O'Connel M, et al. Oxybutynin extended release and tolterodine immediate release. *Clin Drug Invest* 2004;**24**(2):81–8.
4. Guidance for Industry: SUPAC-MR: Modified Release Solid Oral Dosage Forms Scale-Up and Postapproval Changes: Chemistry, Manufacturing, and Controls. In vitro dissolution testing and in vivo bioequivalence documentation. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research; 1997.
5. Wise DL, Klibanov AM, Langer R, Mikos AG, Peppas NA, Trantolo DJ, Wnek GE, Yaszeski MJ, editors. *Handbook of pharmaceutical controlled release technology*. New York, NY: Marcel Dekker, Inc; 2000.
6. Robinson JR, Lee VHL, editors. *Controlled drug delivery: fundamentals and applications*. 2nd ed. New York, NY: Marcel Dekker, Inc; 1987.
7. Kydonieus A. Fundamental concepts of controlled release. In: AF Kydonieus, editor. Controlled release technologies: methods, theory, and applications; CRC Press, Boca Raton, FL 1980, vol. 1, pp. 7.
8. Lee PI, Good WR, editors. *Controlled release technology: pharmaceutical applications, ACS symposium series no. 348*. Washington, DC: American Chemical Society; 1987.
9. Liu Z, Zhou D, Hoag S, Qiu Y. Influence of drug properties and formulation on in vitro drug release and biowaiver regulation of oral extended release dosage forms. *AAPS J* 2016;**18**(2):333–45.
10. Lee, PI. Evolution of oral controlled/modified release dosage forms. In: Proc. CAE/AAPS/CPA/CRS/FIP/APSTJ Conference: Oral Controlled Release Development and Technology, Shanghai; 2007. p. 22–31.
11. Lee PI, Li J-X. Evolution of oral controlled release dosage forms. In: Wen H, Park K, editors. *Oral controlled release formulation design and drug delivery: theory to practice*. Hoboken, NJ: John Wiley & Sons; 2010. p. 21–31.
12. *Recipes for Fifty-Two Ailment*, Mawangdui Medical Manuscript. Dated 168 B.C.
13. Ko Hung, *Handbook of Prescriptions for Urgent Cares*. 281-341 AD.
14. Higuchi T. Rate of release of medicaments from ointment bases containing drugs in suspension. *J Pharm Sci* 1961;**50**:847.
15. Higuchi T. Mechanism of sustained release medication, theoretical analysis of rate of release of solid drugs dispersed in solid matrices. *J Pharm Sci* 1963;**52**:1145.
16. Martin A, Bustamante P, Chun AHC. *Physical pharmacy*. 4th ed. Malvern, PA: Lea & Febiger; 1993. p. 324–55.
17. Crank J. Diffusion in heterogeneous media. In: Crank J, editor. *The mathematics of diffusion*. 2nd ed. Oxford: Claren Don Press; 1986. p. 266–85.
18. Lee PI. Modeling of drug release from matrix systems involving moving boundaries: approximate analytical solutions. *Int J Pharm* 2011;**418**:18–27.
19. Lee P. Diffusional release of a solute from a polymeric matrix—approximate analytical solutions. *J Membr Sci* 1980;**7**:255–75.
20. Foster TP, Parrott EL. Release of highly water-soluble medicinal compounds from inert heterogeneous matrixes. II. Melt. *J Pharm Sci* 1990;**79**:938–42.
21. Ritger PL, Peppas NA. A simple equation for description of solute release I. Fickian and non-Fickian release from non-swellable devices in the form of slabs, spheres, cylinders or discs. *J Control Release* 1987;**5**:23–6.
22. Zhou Y, Chu JS, Zhou T, Wu XY. Modeling of drug release from two-dimensional matrix tablets with anisotropic properties. *Biomaterials* 2005;**26**:945–52.
23. Wu XY, Zhou Y. Studies of diffusional release of a dispersed solute from polymeric matrixes by finite element method. *J Pharm Sci* 2000;**88**(10):1050–7.
24. Zhou Y, Wu XY. Finite element analysis of diffusional drug release from complex matrix systems. I. Complex geometries and composite structures. *J Control Rel* 1997;**49**:277–88.

25. Wu XY, Zhou Y. Finite element analysis of diffusional drug release from complex matrix systems. II. Factors influencing release kinetics. *J Control Rel* 1998;**51**:57–72.
26. Yonezawa Y, Ishida S, Suzuki S, Sunada H. Release from or through a wax matrix system. I. Basic release properties of the wax matrix system. *Chem Pharm Bull* 2001;**49**(11):1448–51.
27. Yonezawa Y, Ishida S, Sunada H. Release from or through a wax matrix system. V. Applicability of the square-root time law equation for release from a wax matrix tablet. *Chem Pharm Bull* 2003;**51**(8):904–8.
28. Yonezawa Y, Ishida S, Sunada H. Release from or through a wax matrix system. VI. Analysis and prediction of the entire release process of the wax matrix tablet. *Chem Pharm Bull* 2005;**53**(8):915–18.
29. Colorcon Inc. http://www.colorcon.com/products-formulation/all-products/polymers-controlled-release/hydrophilic-matrix-tablets.
30. Ford J. Design and evaluation of hydroxypropyl methylcellulose matrix tablets for oral controlled release: a historical perspective. 2014. Hydrophilic Matrix Tablets for Oral Controlled Release. Volume 16 of the series AAPS Advances in the Pharmaceutical Sciences Series. Ed. Timmins P, Pygall SR and Melia CD. Springer, New York, NY, pp. 17–51.
31. Narasimhan B. Accurate models in controlled drug delivery systems. In: Wise DL, Klibanov AM, Langer R, Mikos AG, Peppas NA, Trantolo DJ, Wnek GE, Yaszeski MJ, editors. *Handbook of pharmaceutical controlled release technology*. New York, NY: Marcel Dekker, Inc; 2000. p. 155–81.
32. Colombo P, Santi P, Bettini R, Brazel CS. Drug release from swelling-controlled systems. In: Wise DL, Klibanov AM, Langer R, Mikos AG, Peppas NA, Trantolo DJ, Wnek GE, Yaszeski MJ, editors. *Handbook of pharmaceutical controlled release technology*. New York, NY: Marcel Dekker, Inc; 2000. p. 183–209.
33. Ford JL, Rubinstein MH, McCaul F, Hogan JE, Edgar PJ. Importance of drug type, tablet shape and added diluents on drug release kinetics from hydroxypropylmethylcellulose matrix tablets. *Int J Pharm* 1987;**40**:223–34.
34. Lee PI. Dimensional changes during drug release from a glassy hydrogel matrix. *Polymer* 1983;**24**(Communications):45–7.
35. Lee PI, Peppas NA. Prediction of polymer dissolution in swellable controlled release systems. *J Control Release* 1987;**6**:207–15.
36. Lee PI. Kinetics of drug release from hydrogel matrices. *J Control Release* 1985;**2**:277–88.
37. Lee PI. Controlled drug release from polymeric matrices involving moving boundaries. In: Lewis D, editor. *Controlled release of pesticides and pharmaceuticals*. New York, NY: Plenum Publishing; 1981. p. 39–48.
38. Holley RH, Hopfenberg HB, Stannett V. Anomalous transport of hydrocarbons in polystyrene. *Polym Eng Sci* 1970;**10**:376–82.
39. Ford JL, Mitchell K, Rowe P, Armstrong DJ, Elliott PNC, Rostron C, et al. Mathematical modeling of drug release from hydroxypropyl-methylcellulose matrices: effect of temperature. *Int J Pharm* 1991;**71**:95–104.
40. Peppas NA. Analysis of Fickian and non-Fickian drug release from polymers. *Pharm Acta Helv* 1985;**60**:110–11.
41. Baveja SK, Ranga Rao KV. Sustained release tablet formulation of centperazine. *Int J Pharm* 1986;**31**:169–74.
42. Peppas NA, Sahlin JJ. A simple equation for the description of solute release. III. Coupling of diffusion and relaxation. *Int J Pharm* 1989;**57**:169–72.
43. Ranga Rao KV, Padmalatha DK, Buri P. Influence of molecular size and water solubility of the solute on its release from swelling and erosion controlled polymeric matrices. *J Control Release* 1990;**12**:133–41.
44. Ritger PL, Peppas NA. A simple equation for description of solute release II. Fickian and anomalous release from swellable devices. *J Control Release* 1987;**5**:37–42.
45. Miller-Chou BA, Koenig JL. A review of polymer dissolution. *Prog Polym Sci* 2003;**28**:1223–70.
46. Lee PI. Diffusion-controlled matrix systems. In: Kydonieus A, editor. *Treaties in controlled drug delivery*. New York, NY: Marcel Dekker, Inc; 1992. p. 155–97.
47. Harland RS, Gazzaniga A, Sangalli ME, Colombo P, Peppas NA. Drug/polymer matrix swelling and dissolution. *Pharm Res* 1988;**5**:488–94.
48. Catellani P, Vaona G, Plazzi P, Colombo P. Compressed matrices: formulation and drug release kinetics. *Acta Pharm Technol* 1988;**34**:38–41.
49. Lee PI, Kim CJ. Effect of geometry on swelling front penetration in glassy polymers. *J Membrane Science* 1992;**65**:77–92.
50. Lee PI, Kim C-J. Probing the mechanisms of drug release from hydrogels. *J Control Release* 1991;**16**:229–36.
51. Ju RT, Nixon PR, Patel MV, Tong DM. Drug release from hydrophilic matrices. 2. A mathematical model based on the polymer disentanglement concentration and the diffusion layer. *J Pharm Sci* 1995;**84**(12):1464–77.
52. Ju RT, Nixon PR, Patel MV. Drug release from hydrophilic matrices. 1. New scaling laws for predicting polymer and drug release based on the polymer disentanglement concentration and the diffusion layer. *J Pharm Sci* 1995;**84**(12):1455–63.
53. Liu P, Ju TR, Qiu Y. In: Li X, Jasti BR, editors. *Diffusion-controlled drug delivery systems in design of controlled drug delivery systems*. New York, NY: McGraw-Hill Companies. Inc.; 2006. p. 107–37.
54. Lee PI. Swelling and dissolution kinetics during peptide release from erodible anionic gel beads. *Pharm Res* 1993;**10**(7):980–5.
55. Siepmann J, Peppas NA. Modeling of drug release from delivery systems based on hydroxypropyl methylcellulose (HPMC). *Adv Drug Del Rev* 2001;**48**:139–57.
56. Hildgen P, McMullen J. A new gradient matrix: formulation and characterization. *J Contr Rel* 1995;**34**:263–71.
57. Kim C. Compressed Donut-shaped tablets with zero-order release kinetics. *Pharm Res* 1995;**12**:1045–8.
58. Benkorah A, McMullen J. Biconcave coated, centrally perforated tablets for oral controlled drug delivery. *J Contr Rel* 1994;**32**:155–60.
59. Conte U, Maggi L, Colombo P, Manna A. Multi-layered hydrophilic matrices as constant release devices (Geomatrix™ Systems). *J Contr Rel* 1993;**26**:39–47.
60. Scott D, Hollenbeck R. Design and manufacture of a zero-order sustained-release pellet dosage form through nonuniform drug distribution in a diffusional matrix. *Pharm Res* 1991;**8**:156–61.
61. Brooke D, Washkuhn R. Zero-order drug delivery system: theory and preliminary testing. *J Pharm Sci* 1977;**66**:159–62.
62. Lipper R, Higuchi W. Analysis of theoretical behavior of a proposed zero-order drug delivery system. *J Pharm Sci* 1977;**66**:163–4.
63. Hsieh D, Rhine W, Langer R. Zero-order controlled release polymer matrices for micro and macro-molecules. *J Pharm Sci* 1983;**72**:17–22.
64. Lee IP. Novel approach to zero-order drug delivery via immobilized nonuniform drug distribution in glassy hydrogels. *J Pharm Sci* 1984;**70**:1344–7.
65. Lee IP. Effect of non-uniform initial drug concentration distribution on the kinetics of drug release from glassy hydrogel matrices. *Polymer* 1984;**25**:973–8.
66. Kuu W-Y, Yalkowsky SH. Multiple-hole approach to zero-order release. *J Pharm Sci* 1985;**74**(9):926–33.
67. Kim C. Controlled release from triple layer, donut-shaped tablets with enteric polymers. *AAPS PharmSciTech* 2005;**6**(3):E429–36.

68. Bajpai SK, Bajpai M, Saxena S, Dubey S. Quantitative interpretation of the deviation from 'Zero-Order' kinetics for the release of cyanocobalamin from a starch-based enzymatically degradable hydrogel. *J Macromol Sci, Part A* 2006;**43**(8):1273–7.
69. Wang C-C, Tejwani (Motwani) MR, Roach WJ, Kay JL, Yoo J, Surprenant HL, et al. Development of near zero-order release dosage forms using three-dimensional printing (3-DP™) technology. *Drug Dev Ind Pharm* 2006;**32**(3):367–76.
70. Kshirsagar R, Joshi M, Raichandani Y. Extended release formulation of levetiracetam. US Patent Application No. 20060165796.
71. Delargy AM, Timmins P, Minhom C, Howard JR. The Optimization of a pH independent matrix for the controlled release of drug materials using *in vitro* modeling. *Proc Intern Symp Control Rel Bioact Mater* 1989;**16**:378–9.
72. Ugarkovic S, Trendovska-Serafimovska G, Sapkareva B. Controlled release formulation of verapamil hydrochloride tablets. *Acta Pharma* 1996;**46**:155–7.
73. Badawy SI, Hussain MA. Microenvironmental pH modulation in solid dosage forms. *J Pharm Sci* 2007;**96**(5):948–59.
74. Taniguchi C, Kawabata Y, Wada K, Yamada S, Onoue S. Microenvironmental pH-modification to improve dissolution behavior and oral absorption for drugs with pH-dependent solubility. *Expert Opin Drug Deliv* 2014;**11**(4):505–16.
75. Bassi P, Kaur G. pH modulation: a mechanism to obtain pH-independent drug release. *Expert Opin Drug Deliv* 2010;**7**(7):845–56.
76. Pygall SR, Kujawinski S, Timmins P, Melia CD. Mechanisms of drug release in citrate buffered HPMC matrices. *Int J Pharm* 2009;**370**(1-2):110–20.
77. Williams HD, Ward R, Hardy IJ, Melia CD. The extended release properties of HPMC matrices in the presence of dietary sugars. *J Control Release* 2010;**138**:251–9.
78. Williams HD, Ward R, Hardy IJ, Melia CD. The effect of sucrose and salts in combination on the drug release behaviour of an HPMC matrix. *Eur J Pharm Biopharm* 2010;**76**(3):433–6.
79. Metolose, Water-soluble ethers. Shin-Etsu Chemical Co., Ltd, Tokyo, 2005 (http://www.metolose.ru/files/metolose.pdf).
80. Wells ML, Williams SO, Sanftleben RA, Balik SB, Evans BA. Investigation into the dissolution rate increase on storage of Wellbutrin SR 100 mg tablets. *AAPS PharmSciTech* 2010;**11**(1):113–19.
81. Tatavarti AS, Hoag SW. Microenvironmental pH modulation based release enhancement of a weakly basic drug from hydrophilic matrices. *J Pharm Sci* 2006;**95**(7):1459–68.
82. Tatavarti AS, Mehta KA, Augsburger LL, Hoag SW. Influence of methacrylic and acrylic acid polymers on the release performance of weakly basic drugs from sustained release hydrophilic matrices. *J Pharm Sci* 2004;**93**(9):2319–31.
83. Venkatramana R, Engh K, Qiu Y. Design of pH-independent controlled release matrix tablets for acidic drugs. *Int J Pharm* 2000;**252**(1-2):81–6.
84. Howard JR, Timmins P. Controlled release formulation. United States Patent 4,792,452, December 20, 1988.
85. Zhang G, Pinnamaraju P. Sustained release formulation containing three different types of polymers. United States Patent 5,695,781, December 9, 1997.
86. Rao V, Haslam J, Stella V. Controlled and complete release of a model poorly water-soluble drug, prednisolone, from hydroxypropyl methylcellulose matrix tablets using (SBE)(7m)-beta-cyclodextrin as a solubilizing agent. *J Pharm Sci* 2001;**90**:807–16.
87. Po LW, et al. Characterization of commercially available theophylline sustained- or controlled-release systems: in-vitro drug release profiles. *Int J Pharm* 1990;**66**:111–30.
88. Pillay V, Fassihi R. A novel approach for constant rate delivery of highly soluble bioactives from a simple monolithic system. *J Contr Rel* 2000;**67**:67–78.
89. Hite M, Federici C, Turner S, Fassihi R. Novel design of a self-correcting monolithic controlled-release delivery system for tramadol. *Drug Del Tech* 2003;**3**:48–55.
90. Tanaka N, Imai K, Okimoto K, Ueda S, Tokunaga Y, Ohike A, et al. Development of novel sustained-release system, disintegration-controlled matrix tablet (DCMT) with solid dispersion granules of nilvadipine. *J Control Release* 2005;**108**(2-3):386–95.
91. Abdul S, Poddar SS. A flexible technology for modified release of drugs: multi layered tablets. *J Control Release* 2004;**97**(3):393–405.
92. Qiu Y, Chidambaram N, Flood K. Design and evaluation of layered diffusional matrices for zero-order sustained-release. *J Control Release* 1998;**51**:123–30.
93. Qiu Y, Flood K, Marsh K, Carroll S, Trivedi J, Arneric SP, et al. Design of sustained-release matrix systems for a highly water-soluble compound, ABT-089. *Int J Pharm* 1997;**157**:43–52.
94. Ozeki T, Yuasa H, Okada H. Controlled release of drug via methylcellulose-carboxyvinylpolymer interpolymer complex solid dispersion. *AAPS PharmSciTech* 2005;**6**(2):E231–6.
95. Levina M, Rajabi-Siahboomi AR. The influence of excipients on drug release from hydroxypropyl methylcellulose matrices. *J Pharm Sci* 2004;**93**(11):2746–54.
96. Bajwa GS, Hoebler K, Sammon C, Timmins P, Melia CD. Microstructural imaging of early gel layer formation in HPMC matrices. *J Pharm Sci* 2006;**95**(10):2145–57.
97. Kazarian SG, van der Weerd J. Simultaneous FTIR spectroscopic imaging and visible photography to monitor tablet dissolution and drug release. *Pharm Res* 2008;**25**(4):853–60.
98. Donbrow M, Friedman M. Enhancement of permeability of ethyl cellulose films for drug penetration. *J Pharm Pharmacol* 1975;**27**:633.
99. Donbrow M, Samuelov Y. Zero order drug delivery from double-layered porous films: release rate profiles from ethyl cellulose, hydroxypropyl cellulose and polyethylene glycol mixtures. *J Pharm Pharmacol* 1980;**32**:463.
100. Rowe RC. The effect of the molecular weight of ethyl cellulose on the drug release properties of mixed films of ethyl cellulose and hydroxypropyl methylcellulose. *Int J Pharm* 1986;**29**:37–41.
101. Porter S, Ghebre-Sellassie I. Key factors in the development of modified-release pellets. In: Ghebre-Sellassie I, editor. *Multiparticulate oral drug delivery*. New York, NY: Marcel Dekker; 1994. p. 217–84.
102. Theeuwes F, Higuchi T. Osmotic dispensing device with maximum and minimum sizes for the passageway. US patent 3,916,899, November 4, 1975.
103. McClelland GA, Sutton SC, Engle K, Zentner GM. The solubility-modulated osmotic pump: in vitro/in vivo release of diltiazem hydrochloride. *Pharm Res* 1991;**8**:88–92.
104. Makhija SN, Vavia PR. Controlled porosity osmotic pump-based controlled release systems of pseudoephedrine. I. Cellulose acetate as a semipermeable membrane. *J Control Release* 2003;**89**(1):5–18.
105. Prabakaran D, Singh P, Kanaujia P, Mishra V, Jaganathan KS, Vyas SP. Controlled porosity osmotic pumps of highly aqueous soluble drug containing hydrophilic polymers as release retardants. *Pharm Dev Technol* 2004;**9**(4):435–42.
106. Okimoto K, Tokunaga Y, Ibuki R, Irie T, Uekama K, Rajewski RA, et al. Applicability of (SBE)7m-beta-CD in controlled-porosity osmotic pump tablets (OPTs). *Int J Pharm* 2004;**286**(1-2):81–8.

107. Sotthivirat S, Haslam JL, Stella VJ. Controlled porosity-osmotic pump pellets of a poorly water-soluble drug using sulfobutylether—cyclodextrin, (SBE)7M—CD, as a solubilizing and osmotic agent. *J Pharm Sci* 2007;**96**(9):2364–74.
108. Sotthivirat S, Haslam JL, Lee PI, Rao VM, Stella VJ. Release mechanism of a sparingly water-soluble drug from controlled porosity-osmotic pump pellets using sulfobutylether-β-cyclodextrin as both a solubilizing and osmotic agent. *J Pharm Sci* 2009;**98**(6):1992–2000.
109. Anand V, Kandarapu R, Garg S. Ion-exchange resins: carrying drug delivery forward. *Drug Discov Today* 2001;**6**(17):905–14.
110. Salib NN, El-Menshawy ME, Ismail AA. Polysalt flocculates as a physicochemical approach in the development of controlled-release oral pharmaceuticals. *Pharmazie* 1976;**31**(12):872–4.
111. Bonferoni MC, Rossi S, Ferrari F, Caramella C. Development of oral controlled release tablet formulations based on diltiazem–carrageenan complex. *Pharm Dev Technol* 2004;**9**(2):155–62.
112. Chaudhry NC, Saunders L. Sustained release of drugs from ion exchange resins. *J Pharm Pharmacol* 1956;**38**:975–86.
113. Chaubal MV. Synthetic polymer-based ion exchange resins: excipients & actives. *Drug Deliv Technol* 2003;**3**:6–8.
114. Abdekhodaie MJ, Wu XY. Drug release from ion-exchange microspheres: mathematical modeling and experimental verification. *Biomaterials* 2008;**29**:1654–63.
115. Abdekhodaie M, Wu XY. Drug loading onto ion exchange microspheres: modeling study and experimental verification. *Biomaterials* 2006;**27**:3652–62.
116. Sriwongjanya M, Bodmeier R. Effect of ion exchange resins on the drug release from matrix tablets. *Eur J Pharm Biopharm* 1998;**46**(3):321–7.
117. Klausner EA, Lavy E, Friedman M, Hoffman A. Expandable gastroretentive dosage forms. *J Control Release* 2003;**90**:143–62.
118. Streubel A, Siepmann J, Bodmeier R. Gastroretentive drug delivery systems. *Exp Opin Drug Deliv* 2006;**3**(2):217–33.
119. Kagana L, Lapidotb N, Afarganb M, Kirmayerb D, Moora E, Mardorc Y, et al. Gastroretentive accordion pill: enhancement of riboflavin bioavailability in humans. *J Control Release* 2006;**113**(3):208–15.
120. LeWitt PA, Giladi N, Gurevich T, Shabtai H, Djaldetti R, Roizen N, et al. Accordion pill carbidopa/levodopa (AP-CD/LD) for treatment of advanced PD [abstract]. *Mov Disor* 2014;**29**(Suppl. 1):668.
121. AP—CD/LD: The Accordion Pill™ Carbidopa/Levodopa (http://intecpharma.com/pipleline/ap-cdld/).
122. Pozzi F, Furlani P, Gazzaniga A, Davis SS, Wilding IR. The TIME CLOCK system: a new oral dosage form for fast and complete release of drug after a predetermined lag time. *J Control Release* 1994;**3**:99–108.
123. Ashford M, Fell JT. Targeting drugs to the colon: delivery systems for oral administration. *J Drug Target* 1994;**2**:241–58.
124. Van den Mooter G. Colon drug delivery. *Exp Opin Drug Deliv* 2006 Jan;**3**(1):111–25.
125. Basit AW. Advances in colonic drug delivery. *Drugs* 2005;**65**(14):1991–2007.
126. Chourasia MK, Jain SK. Pharmaceutical approaches to colon targeted drug delivery systems. *J Pharm Pharm Sci* 2003;**6**(1):33–66.
127. Shareef MA, Khar RK, Ahuja A, Ahmad FJ, Raghava S. Colonic drug delivery: an updated review. *AAPS PharmSci* 2003;**5**(2):E17.
128. Philip AK, Philip B. Colon targeted drug delivery systems: a review on primary and novel approaches. *Oman Med J* 2010;**25**(2):79–87.
129. Maroni A, Zema L, Cerea M, Sangalli ME. Oral pulsatile drug delivery systems. *Exp Opin Drug Deliv* 2005;**2**(5):855–71.
130. Arora S, Ali J, Ahuja A, Baboota S, Qureshi J. Pulsatile drug delivery systems: An approach for controlled drug delivery. *Indian J Pharm Sci* 2006;**68**(3):95–300.
131. Till B, Ina O, Rolan B. Pulsatile drug-delivery systems Critical reviews in therapeutic drug carrier systems. *Crit Rev Ther Drug carr Syst* 2001;**18**(5):433–58.
132. Anal Anil K. Time-controlled pulsatile delivery systems for bioactive compounds. *Recent Patents Drug Deliv Formul* 2007;**1**(1):73–7.
133. Jain D, Raturi R, Jain V, Bansal P, Singh R. Recent technologies in pulsatile drug delivery systems. *Biomatter* 2011;**1**(1):57–65.
134. Maroni A, Zema L, Del Curto MD, Loreti G, Gazzaniga A. Oral pulsatile delivery: rationale and chronopharmaceutical formulations. *Int J Pharm* 2010;**398**(1-2):1–8.
135. Eichel HJ, Massmann BD, Cobb JE Jr. Delayed, sustained-release diltiazem pharmaceutical preparation. US Patent 5,529,790.
136. Hendrickson DL, Dimmitt DC, Williams MS, Skultety PF, Baltezor MJ. Diltiazem formulation. US Patent 5,286,497, February 15, 1994.
137. Barnwell SG. Biphasic release formations for lipophilic acids. US Patent 5,391,377, February 21, 1995.
138. Cheung WK, Silber BM, Yacobi A. Pharmacokinetic principles in the design of immediate-release components in sustained-release formulations with zero-order release characteristics. *J Pharm Sci* 1991;**80**(2):142–8.
139. Buttgereit F, Doering G, Schaeffler A, Witte S, Sierakowski S, Gromnica-Ihle E, et al. Efficacy of modified-release versus standard prednisone to reduce duration of morning stiffness of the joints in rheumatoid arthritis (CAPRA-1): a double-blind, randomized controlled trial. *Lancet* 2008;**371**:205–13.
140. Radebaugh GW, Murtha JL, Glinecke R. Oral sustained release acetaminophen formulation and process. US Patent 4,968,509, November 6, 1990.
141. Eichel HJ, Massmann BD. Sustained-release pharmaceutical preparation. US Patent 5,026,559, June 25, 1991.
142. Points to Consider on the Clinical Requirements o Modified Release Products Submitted as a Line Extension of an Existing Marketing Authorization. Committee for Proprietary Medicinal products (CPMP), European Medicines Agency. December, 2003 (http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003124.pdf).
143. Guideline on the pharmacokinetic and clinical evaluation of modified release dosage forms. EMA/CPMP/EWP/280/96 Corr1. Committee for Medicinal Products for Human Use (CHMP), European Medicines Agency, November, 2014.
144. RxKinetics, 2015, (www.rxkinetics.com/antibiotic_pk_pd.html).
145. Ahmed GF, Noreddin A. Application of pharmacokinetics/pharmacodynamics (PK/PD) in designing effective antibiotic treatment regimens. In: Noreddin A, editor. *Readings in advanced pharmacokinetics—theory, methods and applications*. Rijeka: InTech; 2012. p. 111–26.
146. Johan W, Mouton JW, Dudley MN, Cars O, Derendor H, Drusano GL. Standardization of pharmacokinetic/pharmacodynamic (PK/PD) terminology for anti-infective drugs: an update. *J Antimicrob Chemother* 2005;**55**(5):601–7.
147. Nielsen EI, Cars O, Friberg LE. Pharmacokinetic/pharmacodynamic (PK/PD) indices of antibiotics predicted by a semimechanistic PKPD model: a step toward model-based dose optimization. *Antimicrob Agents Chemother* 2011;**55**(10):4619–30.
148. Liu P, Fang AF, LaBadie RR, Crownover PH, Arguedas AG. Comparison of azithromycin pharmacokinetics following single oral doses of extended-release and immediate-release formulations in children with acute otitis media. *Antimicrob Agents Chemother* 2011;**55**(11):5022–6.

149. Mager DE, Wyska E, Jusko WJ. Diversity of mechanism-based pharmacodynamic models. *Drug Metab Dis* 2003;**31**(5):510–19.
150. Breimer DD. An integrated pharmacokinetic and pharmacodynamic approach to controlled drug delivery. *J Drug Target* 1996;**3**:411–15.
151. Hoffman A. Pharmacodynamic aspects of sustained release preparations. *Adv Drug Deliv Rev* 1998;**33**:185–99.
152. Shugarts S, Benet LZ. The role of transporters in the pharmacokinetics of orally administered drugs. *Pharm Res* 2009;**26**(9):2039–54.
153. Sutton SC. Companion animal physiology and dosage form performance. *Adv Drug Deliv Rev* 2004;**56**(10):1383–98.
154. Bauer KH, Lehmann K, Osterwald HP, Rothgang G. *Environmental conditions in the digestive tract and their influence on the dosage form. Coated pharmaceutical dosage forms, fundamentals, manufacturing techniques, biopharmaceutical aspects, test methods and raw materials*. Birkenwaldstr: Medpharm GmbH Scientific Publishers; 1998. p. 126–30.
155. Abrahamssona B, Alpstenb M, Bakec B, Jonssona UE, Eriksson-Lepkowskaa M, Larsson A. Drug absorption from nifedipine hydrophilic matrix extended-release (ER) tablet-comparison with an osmotic pump tablet and effect of food. *J Control Release* 1998;**52**(3):301–10.
156. Qiu Y, Cheskin H, Engh K, Poska R. Once-a-day controlled-release dosage form of divalproex sodium I: formulation design and in vitro/in vivo investigations. *J Pharm Sci* 2003;**92**(6):1166–73.
157. Vinks AA, Walson PD. Chapter 4. Pharmacokinetics I: Developmental principles. In: Martin A, Scahill L, Charney DS, Leckman JF, editors. *Pediatric psychopharmacology: principles and practice*. New York, NY: Oxford University Press; 2003. p. 49.
158. Sutton S, Evans L, Fortner J, McCarthy J, Sweeney K. Dog colonoscopy model for predicting human colon absorption. *Pharm Res* 2006;**23**(7):1554–63.
159. Wilson CG, Wahsington N. Gamma scintigraphy in the analysis of the behaviors of controlled release systems. In: Wise DL, Klibanov AM, Langer R, Mikos AG, Peppas NA, Trantolo DJ, Wnek GE, Yaszeski MJ, editors. *Handbook of pharmaceutical controlled release technology*. New York, NY: Marcel Dekker, Inc; 2000. p. 551–66.
160. Parasrampuria DA, Kanamaru T, Connor A, Wilding I, Ogata K, Shimoto Y. Kunitada. Evaluation of regional gastrointestinal absorption of edoxaban using the Enterion capsule. *J Clin Pharmacol* 2015;**55**(11):1286–92.
161. Stevens HNE, Speakman M. Behaviour and transit of tamsulosin Oral Controlled Absorption System in the gastrointestinal tract. *Curr Med Res Opin* 2006;**22**(12):2323–8.
162. Söderlind E, Abrahamsson B, Erlandsson F, Wanke C, Iordanov V, von Corswant C. Validation of the IntelliCap® system as a tool to evaluate extended release profiles in human GI tract using metoprolol as model drug. *J Control Release* 2015;**217**(10):300–7.
163. Koziolek M, Grimm M, Becker D, Iordanov V, Zou H, Shimizu J, et al. Investigation of pH and temperature profiles in the gi tract of fasted human subjects using the Intellicap® system. *J Pharm Sci* 2015;**104**:2855–63.
164. Gualtieri, et al. Clinical studies of methylphenidate serum levels in children and adults. *J Am Acad Child Psyc* 1982;**21**:19–26.
165. Pentikis HS, Gonzalez MA. Effect of formulation on methylphenidate release patterns: clinical implications. *Am J Drug Deliv* 2005;**3**(1):47–54.
166. Gonzalez MA, Pentikis HS, Anderl N, Benedict MF, DeCory HH, Dirksen SJH, et al. Methylphenidate bioavailability from two extended-release formulations. *Int J Clin Pharmacol Ther* 2002;**40**:175–84.
167. Sonuga-Barke EJS, Van Lier P, Swanson JM, Coghill D, Wigal S, Vandenberghe M, et al. Heterogeneity in the pharmacodynamics of two long-acting methylphenidate formulations for children with attention deficit/hyperactivity disorder: a growth mixture modeling analysis. *Eur Child Adolesc Psychiatry* 2007. Published online: 10 December
168. Markowitz JS, Straughn AB, Patrick KS, DeVane CL, Pestreich L, Lee J, et al. Pharmacokinetics of methylphenidate after oral administration of two modified-release formulations in healthy adults. *Clin Pharmacokinet* 2003;**42**(4):393–401.
169. Wang Y, Lee L, Somma R, Thompson G, Bakhtiar R, Lee J, et al. In vitro dissolution and in vivo oral absorption of methylphenidate from a bimodal release formulation in healthy volunteers. *Biopharm Drug Dispos* 2004;**25**(2):91–8.
170. BIAXIN® XL Filmtab® (clarithromycin extended-release tablets). *Physician's desk reference*. 59th ed Montvale, NJ: Thompson PDR; 2005. p. 408–17.
171. Saigala N, Babootaa S, Ahujaa A, Ali J. Multiple-pulse drug delivery systems: setting a new paradigm for infectious disease therapy. *Exp Opin Drug Deliv* 2009;**6**(4):441–52.
172. Dmochowski R, Staskin D, Kell S. Oxybutynin chloride: alterations in drug delivery and improved therapeutic index. *Exp Opin Pharmacother* 2002;**3**(4):443–54.
173. Michel MC, Benefit-Risk A. Assessment of extended-release oxybutynin. *Drug Safety* 2002;**25**(12):867–76.
174. Dmochowski RR, Staskin DR. Advances in drug delivery: improved bioavailability and drug effect. *Curr Urol Reports* 2002;**3**(6):1527–2737.
175. Good WR, Lee PI. Membrane-controlled reservoir drug delivery systems. In: Langer RS, Wise DL, editors. *Medical applications of controlled release*, vol. I. Boca Raton, FL: CRC Press, Inc.; 1984. p. 1–39.

CHAPTER

20

Product and Process Development of Solid Oral Dosage Forms

Y. Qiu[1], *X. He*[2], *L. Zhu*[1,3] *and B. Chen*[1]

[1]AbbVie, Inc., North Chicago, IL, United States [2]Boehringer Ingelheim, Ridgefield, CT, United States
[3]Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, PR China

20.1 INTRODUCTION

A pharmaceutical dosage form development program typically includes preformulation studies, analytical method development and validation, design, development, scale-up and optimization of formulation and manufacturing process, in vivo bioavailability and stability studies. This chapter focuses on transforming technology-based dosage form design into commercial drug products that meet patient requirements. The technology transformation process requires design and development of effective formulation and manufacturing process that assure consistent product quality and performance. In addition, it should also take into consideration regulatory requirements, manufacturability, process capability, robustness, efficiency and control, equipment, facility, capacity, development cost and time, environmental impact, and market dynamics, among others. For solid oral products, these considerations often have strong implications in developing a reliable dosage form and robust process to assure consistent product quality, manufacturability, and productivity, or in developing new formulation or manufacturing technologies for out-licensing.

A solid dosage form is a complex multicomponent heterogeneous solid system that is typically subject to mechanical stresses during its production. The majority of the components are organic compounds with a very wide range of physical, chemical, and mechanical properties, as opposed to the inorganic components, such as metals and solid solution alloys used in the high-tech and aerospace industries. Combining these different solid materials via pharmaceutical processing to form a product is usually accompanied by changes in density and volumes of coexisting solid phases, which can produce large stresses and strains, alter thermodynamic properties of the materials and affect mechanical and chemical equilibrium within and between solid phases and ultimately the structure of the material. As a result, product quality, performance, and/or processability usually depend on chemical, physical or mechanical properties, and compatibility of these materials. The complex interactions involved are difficult to understand as engineering and scientific models are unavailable to accurately determine or predict structures and processing behavior of physically combined solid powders containing multiple sources of stress, including thermal, mechanical, and gravitational. Moreover, the multistage batch process commonly used in pharmaceutical manufacturing makes it more difficult to determine the process controls and outputs due to the presence of a multitude of interacting material, formulation and process variables. For instance, the governing equations of the dynamic flow, stochastic mixing and bonding of solid particles with diverse morphology, and the density and surface properties in the presence or absence of a liquid are undefined or poorly understood in the context of high-shear wet granulation, drying, or blending, limiting the ability to reliably model these unit operations. Furthermore, a product's ultimate test is its intended clinical in vivo performance in patients. However, universal in vitro biopharmaceutics tools or models are unattainable for evaluating and predicting possible impact of dosage form attributes and process parameters on the bioavailability performance of various drug products due to complexity of

DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00020-0

biological systems. Hence, developing pharmaceutical products and ensuring consistent quality are both complex and challenging.

The present state of the art of linking product quality to the desired biological performance is to consider and test in vivo performance in the design and development phases of the drug product and process using rational approaches and predictive models where possible on a product-by-product basis. Therefore, for a product developed under a set of conditions, even a small incremental change/adjustment of the composition, processing steps or conditions, or equipment design, often requires an investment of extensive resource and time to generate experimental data to demonstrate that the product quality remains unchanged. This is partly caused by a lack of understanding of, and inability to predict, the possible impact of such change because of the complexity mentioned earlier of solid dosage forms. It is also affected by the conservative philosophy and policy adopted by both the sponsors and regulators since the pharmaceutical industry is one of the most regulated industries that generally requires very exacting quality standards. However, these challenges should not deter scientists and engineers from applying principles of multidisciplinary science and a systematic approach in developing solid dosage forms.

20.2 DEVELOPMENT OF SOLID DOSAGE FORMS

Because of the complexity of solid dosage forms and limitations and challenges in applying First-Principles approach in the pharmaceutical industry, the strategies, and approaches that have been, and continue to be, utilized in product and process development vary significantly from company to company, and even across project teams within the same organization. Some groups tend to rely more heavily on empirical studies and prior knowledge that can get the job done, albeit often inefficiently. Some work with past experiences or "gut feeling" that may or may not be valid or applicable. Others take a more rational approach by systematically applying applicable scientific and engineering principles to development. In general, these practices can be divided into three broad categories summarized in Table 20.1.

Today, it is widely recognized and accepted that adopting a rational development approach with increased level of scientific understanding can result in many benefits for patients, industry, and regulatory agencies. Advantages can include a higher level of assurance of product quality and performance, more robust manufacturing process and process control, increased cost savings and improved efficiency through minimized product rejects or recalls, reduced regulatory burden and enhanced opportunities for the first cycle product approval. It also helps to identify risks, contributes to the development of risk management strategy, streamlines postapproval changes, and offers opportunities for postapproval continuous improvement. In fact, this is the approach that the Food and Drug Administration (FDA) has been advocating as part of its ongoing quality initiative, "cGMPs for the 21st Century," first announced in 2002[1] to enhance and modernize pharmaceutical product quality and manufacturing. Over the past decade, the FDA has collaborated with the International Conference on Harmonization (ICH) and the industry in defining a shared vision, key concepts and general framework of the initiative. Among the critical concepts is pharmaceutical quality by design (QbD).[2] It is a systematic scientific approach to product and process design and development to ensure consistent product quality through understanding and controlling material, formulation, and manufacturing variables. In 2006, the principles of QbD were incorporated in the ICH Q8 (Pharmaceutical Development), Q9 (Quality Risk Management), and Q10 (Quality Risk Management)[3] guidance which has since been updated. In 2007, Office of Generic Drugs (OGD) of FDA officially started implementation of a question-based review (QbR) developed for its chemistry, manufacturing, and control (CMC) evaluation of the abbreviated new drug applications (ANDAs). The QbR is a science and risk-based quality assessment system that incorporates key concepts and elements of QbD. It requires applicants to use the common technical document (CTD) framework to address a series of prescriptive questions by applying QbD principles.[4]

In November 2014, the FDA issued a Manual of Policies and Procedures (MAPP 5015.10) that covers chemistry review of QbR submissions of drug and product applications (NDAs, ANDAs, and DMFs). More recently, FDA's quality initiative was further enhanced by forming a new office, the Office of Pharmaceutical Quality (OPQ) that streamlines processes of monitoring quality throughout the product lifecycle for both innovator and generic drug products. To assure consistent product quality and prevent quality failures that may potentially lead to drug recalls or shortages, product and process understanding is not only expected during the design and development stage but also in the postapproval improvements and commercial manufacturing stage. Specifically, the product and process should be designed to meet product critical quality attributes (CQAs) consistently. Product standards should be captured in scientifically sound and clinically relevant CQAs and specifications; the impact of starting materials and process parameters on product quality and manufacturing should

TABLE 20.1 Comparison of Approach to Solid Product Development

Approach	Characteristics	Likely outcome
Trial-and-error	Paying limited or inadequate attention to properties of raw materials and composition, and relevant scientific principles Trying out various experiments or hypotheses in various directions until a desired outcome is obtained through testing with certain degree of reliability Primarily relying on past experience and intuition: for example, copying "recipe" from one product to another; Using "shotgun" approach	Disconnect between results and underlying mechanism or principle Data generated can be overwhelming or sometimes misleading due to experiments with confounding variables Considerable demand for time and resources, resulting in poor efficiency Inconsistent or nonrobust product and process that can lead to quality failures during development and/or commercial manufacturing
Semiempirical	Using best guessed "trial-and-error" approach and/or observational studies based on prior knowledge and statistical principles with little or limited mechanistic understanding Combining experience with analysis of abundant data to identify trends or build empirical or semiempirical relationships	Results can be practically useful, but may be unreliable or misleading in certain cases due to improper design, limitations of experiments and/or lack of a thorough scrutiny of data and various formulation/process variables involved Higher demand for time and resources for data generation Lower level of assurance of consistent product quality and reliable process
Rational	Characterizing raw materials, formulation, process, and possible interactions Integrating formulation/process design and development by utilizing multidisciplinary scientific and engineering principles and applying a systematic approach Using knowledge-based, probabilistic, and best-guess approach appropriately through combining prior knowledge and data with basic understanding in experimental design	Enhanced understanding of how material properties, process variables, and product attributes relate to product performance, as well as the interplays between biological system and the drug substance or dosage form Increased efficiency and streamlined continual improvement Higher level of assurance for consistent quality of drug product and commercial manufacturing; minimized product rejects or recalls

be understood; critical sources of material and process variation should be identified and controlled; and the manufacturing should be continually monitored with quantitative quality metrics and updated to allow for consistent quality over time. To achieve these objectives, the adoption of the principles of rational product development is essential.

20.2.1 Rational development approach

Product realization process (ISO 9001:2008 Section 7) typically involves product identification, concept development, product, and process design and development including scale-up of the manufacturing process, and commercial production. Rational development of product and process requires integrated consideration of three essential areas: (1) preformulation studies, (2) formulation and process investigation, and (3) in vivo performance evaluation to develop a high quality dosage form with predefined quality attributes and drug delivery performance. The key to attaining a consistently high quality product and robust manufacturing process is to adopt proactively a systematic approach through applying multiple sciences and engineering principles, such as pharmaceutical sciences, chemistry, biological sciences, engineering, mathematics, and statistics as shown in Fig. 20.1.

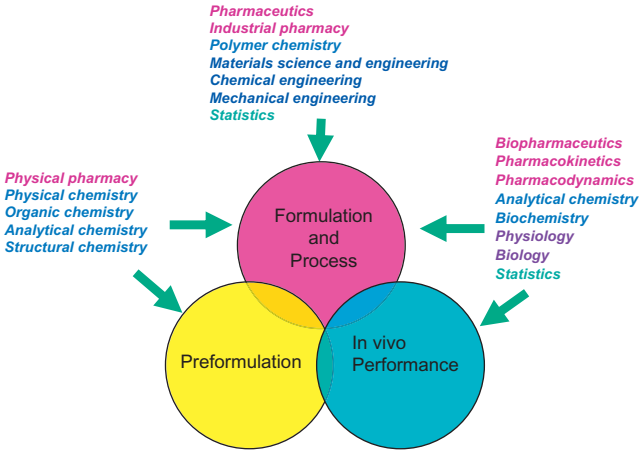


FIGURE 20.1 Multidisciplinary principles important to rational product development.

Development of a solid dosage form starts with defining a quality target product profile (QTPP) based on the clinical and patient needs. In addition to the desired in vivo performance, other important criteria include strength, type, size, drug loading, quality and performance attributes of the dosage form, as well as manufacturing and commercial requirements. The second step is to characterize properties of the active pharmaceutical ingredient (API) and excipients, including their physicochemical, biopharmaceutical,

mechanical properties, and compatibility, which has been discussed in detail in the relevant chapters of this book. Once the quality characteristics of the drug substance and excipients are determined, formulation and process feasibility and options should be evaluated. There are often more than one processing technologies that can be used to produce a solid dosage form with predefined product quality and performance attributes. The best possible manufacturing approach should be selected based on technical, economical, and practical considerations, including compatibility with the components, maturity of technologies, time-to-market, development cost, process complexity, scalability, manufacturability, robustness, commercial production efficiency, and cost along with available equipment, facility, capacity, and footprint. One of the most critical elements in developing a solid dosage form is to ensure a connection between the design of the formulation and manufacturing process. In addition to API, a formulation typically consists of essential excipients called for by a specific dosage form design, such as filler, binder, disintegrant, lubricant, release modifier, physical or chemical stabilizer, processing aid, etc. Selection of these components and a particular process train requires an integrated consideration of material properties, functionality and processing characteristics in connection with properties of a specific API, unit operations and the interactions between them. Such a concurrent development approach not only can result in quality products with consistent performance and robust process, but also help ensure optimum manufacturability and productivity, minimize or avoid potential material-related processing issues or process-related product quality or performance problems.

20.2.2 Integrated formulation and process design

The realization process of pharmaceutical products has undergone a number of significant changes over the last decade. Many of them are brought about, in large part, by the FDA's quality and patient-focused initiatives[1] and increased globalization. A successful dosage form development program requires specific knowledge at three critical stages: (1) designing product and process based on clearly defined requirements, (2) demonstrating a product design that meets product requirements, and (3) producing a reliable product reproducibly within established quality, cost, resource, and schedule targets. To create and efficiently manufacture reliable and consistent high-quality products, it is crucial to use an integrated concurrent approach in conceptualizing, designing, and developing products. This is done by incorporating a broad set of requirements, objectives, and constraints of the product life cycle, including patient needs, product quality, reliability, manufacturing, materials, cost, supply, environment, and safety aspects. In developing pharmaceutical products, drug properties, required dose, and mode of drug delivery typically dictate the choice of the dosage form and determine the overall approach to formulation and process design. Hence, understanding the characteristics and natural variability of formulation components, the process characteristics, and their interplay is the foundation for product development and commercial manufacturing. Specifically, knowledge of API physicochemical, mechanical, and biopharmaceutical properties is essential to the concurrent product and process design because it provides a basis for optimizing product quality and performance and ultimately maximizes the chances of developing stable, safe, effective, and commercially viable dosage forms of consistent quality. In parallel, understanding of the excipient properties and control are often equally important.

20.2.2.1 Material property consideration

For solid oral dosage forms, API properties often have a significant influence on the product's CQAs including purity, uniformity, drug release, and stability, as well as manufacturability. Therefore, it is important to take into account API material properties when designing a drug product and its manufacturing process. The basic concepts and techniques of physicochemical, biopharmaceutical, and mechanical characterization useful in the rational product development have been discussed in the relevant chapters of this book. In this section, we will discuss in details (1) how API physicochemical, powder/bulk, and biopharmaceutical properties can affect drug product quality and performance with a focus on dissolution/bioavailability, stability, and manufacturability and (2) what options are available to overcome challenges posed by such properties? In addition, we will also discuss how to practically integrate API material properties into the development of a fit-for-purpose formulation during various phases of clinical trials in order to balance risk, cost, speed, and quality of new drug development.

20.2.2.1.1 Physicochemical properties

API physicochemical properties can have a significant impact on stability, manufacturability, and in vivo performance of drug products. Therefore, designing formulation and manufacturing process should always start with understanding properties such as solubility, permeability, dissolution rate, ionization constant/pKa, melting point, polymorphism, and crystallinity.

The aqueous solubility of a drug molecule is one of the most critical factors that influence the rate of drug absorption after oral administration. The fundamental aspects of solubility can be found in chapter "Solubility of Pharmaceutical Solids." In general, high aqueous solubility is desired for maximizing the driving force in intestinal drug absorption. The majority of pharmaceutical compounds are absorbed via a passive diffusion process in the small intestine. For drugs that have high permeability, transport across the intestinal membrane is usually fast. However, low aqueous solubility or slow dissolution rate may limit the rate and extent of drug absorption. To enhance oral absorption of poorly soluble drugs, different formulation approaches have been developed and will be described in detail later. Permeability is another important factor that influences the rate and extent of drug absorption following oral administration. There are essentially three mechanisms by which a drug molecule may cross the intestinal membrane: passive diffusion, uptake transporter-mediated transport, and vesicular transport. These are discussed in details in chapter "Oral Absorption Basics: Pathways and Physicochemical and Biological Factors Affecting Absorption." Permeability is a composite of diffusion coefficient, partition coefficient, and membrane thickness. There are practically no formulation options available to effectively and consistently increase local permeability of poorly permeable drugs (such as BCS Class III/IV compounds) in the gastrointestinal (GI) tract. Certain excipients may alter membrane permeability in the in vitro setting, however, their use and effectiveness has not been proven in vivo mainly due to limited quantity present in a dosage form, need for synchronized transit with the API and effect of dilution in the dynamic environment of the intestinal tract. If a local enhancer is found to be effective, it will need first to address potential safety concerns before its regulatory clearance. Dissolution rate (see chapter: "Fundamental of Diffusion and Dissolution") is proportional to both solubility and surface area of API powder. Dissolution rate and free drug concentration in solution may be altered by formulation design to achieve the intended therapeutic exposure. For example, API particle size reduction, use of high energy solid phase, and/or selection of soluble excipients may be used to increase dissolution/solubility rate of less soluble drugs. In other applications, a modified release (MR) dosage form may be designed to reduce dosing frequency and/or achieve more prolonged and consistent exposure with the maintenance of the blood concentration within a specific range. These formulation technologies are described elsewhere in this book.

The majority of small molecules are either weak acids or bases. The ionization (or dissociation) constant pKa has a significant influence on solubility, permeability, dissolution rate, and stability. The total aqueous solubility of a weak acid or base drug is the sum of the concentrations of ionized and unionized species. Therefore, it is heavily influenced by both pKa and pH, which, in turn, could affect drug absorption in the GI tract, where the pH is known to be highly variable. For example, a weak acid with a pKa of 4 has low solubility in the stomach (pH 1−3), but the solubility dramatically increases as the drug enters small intestine where pH ranges from 5.5 to 7. Using a soluble salt helps rapid dissolution of the drug. However, one area of concern of using a soluble salt is that it may precipitate in vivo with abrupt pH changes. For instance, a soluble salt of a weak base with a pKa of 8 could rapidly dissolve in the stomach, but as the drug enters the small intestine, precipitation of the free base could occur. Another important physicochemical property is the melting point, which can be viewed as a measure of "energy" required to overcome attractive forces that "hold a crystal lattice" together.

Crystal lattices with greater molecular symmetry and the presence of hydrogen bonds significantly increase intermolecular interactions in the solid state, resulting in higher melting point. From a product development point of view, melting point is useful information because compounds with a higher melting point tend to have a lower solubility and hence slower dissolution rate. Melting point may also be used to assess and quantify the presence of impurities and the interconversion of solid forms.[5] From a manufacturing perspective, melting points below about 60°C are potentially problematic. Temperatures in conventional solid manufacturing such as fluid bed drying, tablet compression, and coating can exceed 50−60°C. During the milling process, hot spots in the milling chamber may have much higher temperatures. On the other hand, an API (or excipient) with a high melting temperature can be challenging for using melt processes to perform granulation or form molecular dispersion in polymers (eg, melt extrusion or congealing). Thus, the selection of manufacturing process should take into consideration the melting point of both API and excipients.

Polymorphism (see chapter: "Crystalline and Amorphous Solids") can have profound implications in formulation development because polymorphs may exhibit significantly different solubility, dissolution rate, compactibility, hygroscopicity, physical, and chemical stability. A metastable form often has higher solubility and dissolution rate, which may lead to higher oral bioavailability, such as chloramphenicol palmitate.[6] Although the use of a more soluble polymorph may provide higher bioavailability, it is important to keep in mind that a metastable polymorph (higher energy form) tends to convert to a

thermodynamically more stable form over time. The conversion between solid forms is affected by multiple factors, such as solvent, relative humidity, temperature and mechanical agitation. Conversion from a metastable form to a stable form could reduce dissolution rate and lead to inconsistent product quality. Therefore it is often preferred to select the thermodynamically most stable polymorph for development. It is also important to note that polymorph conversion may also occur during manufacturing processing (such as milling, wet or melt granulation, drying) even when a stable crystal form is chosen as a starting material for development.[7] Therefore, one should assess the risk of process induced transformation for any API with multiple solid forms.

Over the past decade, higher apparent solubility and faster dissolution rate of amorphous solids have increasingly been utilized to enhance delivery of insoluble APIs. The basic principles and formulation aspects of utilizing amorphous solids are discussed in details in chapters "Crystalline and Amorphous Solids", "Oral Formulations for Preclinical Studies: Principle, Design, and Development Considerations" and "Rational Design for Amorphous Solid Dispersions." For this type of product, crystallinity is a critical factor to consider during formulation and process development. The degree of crystallinity refers to the ratio of crystalline to amorphous content in a powder. Because the amorphous solid is thermodynamically unstable, it will eventually crystallize to a more stable, and potentially less soluble solid form when given sufficient time and under the right conditions. Generally, the rate of crystallization depends on many factors, such as temperature, relative humidity, processing history, and excipients, etc. To ensure consistent product quality and performance, it is essential to control batch to batch variation in amorphous content with no or minimized crystallinity. In practice, polymers are usually used in the formulation to inhibit crystallization within the desired product shelf-life. It is a rule of thumb that an amorphous solid dispersion (ASD) has reasonable physical stability when it is stored 50°C below its glass transition temperature. Hence, selection of the polymers, manufacturing process, packaging, and storage condition needs to be based on product and process understanding.

20.2.2.1.2 Powder and bulk properties

API powder properties can have an impact on bioavailability, stability, and processing of solid oral dosage forms. In particular, powder properties that include particle size distribution, particle shape, density, porosity, surface area, cohesiveness, and hygroscopicity may affect powder bulk handling properties such as compaction and flow characteristics which may, in turn, impact manufacturability of tablet dosage forms, particularly for formulation containing high percentage of API.[8] For example, to obtain consistent tablet weight and content uniformity during tablet compression on a high speed rotary tablet press, sufficient flow is necessary for discharging a formulation blend from the hopper to the feed frame and distributing it into a die cavity. In addition, an adequate compaction property of the powder blend is also required to form a strong and physically stable compact. Hence, understanding and controlling lot-to-lot variability of key API bulk properties are essential to ensure quality product can be produced consistently.

API particle size is known to affect drug dissolution of poorly soluble drugs. For a low dose (typically 10 mg dose strength or below) direct compression formulation, a small particle size is often needed to overcome manufacturing challenges related to achieving acceptable content uniformity of the API in the formulated blend.[9,10] Particle size, shape and surface properties of a solid also affect powder flow properties. APIs that are needle-like crystals often have poor filtration and flow properties causing issues with both the API isolation and drug product manufacture. Milling of long needles can enhance flow and improve content uniformity; however, excessively small particles tend to be cohesive, which can exacerbate flow problems, resulting in large tablet weight variation during compression for formulations with medium or high drug loading. For instance, the flow of small particles (eg, less than 10 μm) through an orifice is usually restricted because the cohesive forces between the particles are of the same magnitude as gravitational forces.[11] In addition, particle size may also influence compaction properties of pharmaceutical powders.[12] Furthermore, it should be noted that certain physical properties of starting materials can be significantly altered after processing. The bulk properties of a formulation blend at the end of the manufacturing process may be different from the additive properties of its constituents at the beginning of the process. Therefore, it is important to characterize API, excipients as well as formulation blends to gain a better understanding of how powder properties may affect content uniformity, dissolution, and manufacturability.

The density of a solid can affect flow, mixing, and compressibility. Since pharmaceutical particles are often of different sizes and typically have a morphology containing microscopic cracks, internal pores, and capillary spaces, three types of density are used to characterize pharmaceutical powders—true density, bulk density, and tapped density. True density is defined as the density of the material itself, exclusive of any voids or interparticle pores larger than molecular or atomic dimensions in the crystal lattice. Bulk density, which takes into account of macroscopic

interparticle space, is defined as powder mass divided by its bulk volume without any tapping. Tapped density is measured after a powder sample is subjected to mechanically tapping. True density is needed to calculate solid fraction of a compact, which is needed to analyze compactibility and compressibility of powder blends. Because many types of product processing equipment have a fixed volume, bulk density is often useful in calculating the batch size. Bulk density and tapped density can also be used to calculate compressibility index, an empirical tool to assess flow properties. A free flowing powder usually has a low compressibility index as its bulk density is close to the tapped density because the interparticle forces are not as significant as those of the poor flowing powder. As a general rule of thumb, a compressibility index greater than 30% indicates poor powder flow. Porosity not only affects compactability, but also density. It can be estimated from powder density and is an important attribute of pharmaceutical granules.

Granule porosity is typically determined by mercury porosimetry where mercury can fill inter- and intraparticle voids under pressure except the internal pores of the particles. Granules with a low porosity will be harder to compress than the same granules with a high porosity. Porosity may also affect product dissolution rate. It is believed that porous granules dissolve faster than dense granules since the pores allow water to penetrate granules more readily. Processing conditions can have a strong influence on the porosity of a powder. For example, fluid-bed granulation (FBG) is known to produce smaller and more porous granules than high shear wet granulation. Hence, this process is sometimes preferred for certain compounds/formulations that require enhanced dissolution or compressibility.

Hygroscopicity is a material's ability to adsorb and retain water molecules from the surrounding environment. Many drug substances and excipients, particularly water-soluble salts or amorphous API, readily adsorb atmospheric moisture, which can influence physical and chemical stability, flowability, and compatibility of drug products. For example, moisture facilitates conversion between anhydrous and hydrate forms, with potential adverse effects on the solubility and the physical stability of a solid. Some deliquescent materials can adsorb sufficient water to completely dissolve once a critical equilibrium atmospheric humidity (RHCE) is exceeded. As for chemical stability, there is a large body of evidence showing that the adsorbed water can facilitate solid-state reactions through increasing the molecular mobility of the reactants, or itself can participate in the hydrolysis as a reacting species.[13] For moisture sensitive compounds such as aspirin, hydrolysis can occur during the aqueous film coating process of tablets.[14] For moisture labile drugs, it is clearly important to control the amount of water/water activity to which API is exposed during manufacturing and storage.

In certain cases, moisture can have a positive impact on a drug product. For example, it may improve powder flow, and uniformity of the bulk density, as water helps dissipate powder's triboelectrostatic charge. Appropriate amounts of moisture may also facilitate compaction through increased bonding caused by enhanced plastic deformation.[15] Therefore, understanding hygroscopic properties of formulation components is essential in making rational decisions related to formulation, process, packaging, storage, shipping, and handling.

Understanding the mechanical and flow properties of APIs is important in the successful design of formulation and manufacturing process, especially for high drug load formulations where API properties dominate the bulk properties of the formulation. Mechanical properties of a material describe the relationship between applied force and the resulting deformation and are hence critical to understanding powder processing for tableting. In manufacturing tablet dosage forms, the compaction properties of a pharmaceutical powder are typically influenced by many factors, such as particle size, shape, crystal lattice arrangement, interparticle bonding, porosity, surface charge and moisture. A material can be characterized by its compressibility and compactibility. While compressibility is the ability of the powder to deform under pressure, compactibility is the ability to form mechanically strong compacts. Various methods have been used to evaluate the compaction properties with different advantages and limitations. The most common method is to produce a coherent compact by applying a compressive pressure (stress) to a powder, and analyze its tensile strength/hardness. Compaction characteristics of a material or blend including deformation mechanism, yield pressure, Young's modulus, tabletability, compactibility, and compressibility, etc., are usually obtained by varying the rate and extent of the applied stress (compression profile). Similar to compaction properties, flow properties are also influenced by many material properties including particle size and size distribution, particle shape, interparticle interaction, moisture, consolidation stress applied to powder. Characterization of API flow properties is essential in product and process development, as well as selection of solid handling equipment, to eliminate flow problems (eg, flow obstructions, segregation, irregular flow, and flooding) during processing. There are a variety of methods for characterizing flow properties: angle of repose, compressibility index, Hausner ratio, and flow through an orifice are often determined

to provide a rank-order estimation of powder flow properties. More sophisticated instruments, such as shear cells, can be used to characterize powder flow for quantitative performance predictions and engineering design calculations, such as estimation of the hopper angles required for mass flow. Details of powder compaction and flow characterization can be found in chapters "Particle, Powder, and Compact Characterization" and "Process Development, Optimization, and Scale-Up: Providing Reliable Powder Flow and Product Uniformity."

In developing solid drug products, understanding API powder and bulk properties, and the ability to control such properties, is critical for anticipating and preventing potential manufacturing issues, particularly for high and low drug load formulations. For drug substances with undesirable bulk properties, rational formulation and manufacturing process design, and/or API particle engineering can be utilized to overcome processing challenges. In the formulation design, suitable excipients can be selected to accommodate poor compaction and flow properties of the API. As a general rule of thumb, plastically deforming excipients with excellent compactibility should be used if the API is brittle or elastic. For example, microcrystalline cellulose is often used as a compression aid to accommodate poor API compaction properties. For API with poor flow properties, excipients such as glidant or direct compression grade fillers can be used to enhance flow properties of the formulation. Flow and compaction properties of a powder blend can also be improved by selecting an appropriate manufacturing process such as wet, melt, or dry granulation. Granulation technology may also be used to improve API content uniformity for low dose formulation (eg, the drug loading is $<5\%$ w/w). For API that has acceptable bulk properties, direct compression process may be preferred due to its simplicity and low cost. However, direct compression process is more susceptible to lot to lot variation of API and excipient material properties than the granulation processes. Selection of the manufacturing process also strongly depends on the type of dosage forms to be manufactured. For compounds that require enabling formulation technology (such as ASDs) to enhance oral bioavailability, a hot melt extrusion, or spray drying process may be required to generate amorphous API.

Finally, API powder properties may be modified by particle engineering technologies during the API isolation process because particle morphology, surface properties and bulk density are known to be dependent on crystallization process parameters such as crystallization solvents, cooling rate, particle properties of seeds, and seeding amount. In addition, API bulk properties may be further modified by altering the dryer design and drying process. As discussed previously, a needle-shape API can result in potential filtration, flow, and compaction challenges. Through API particle engineering it might be possible to design the API with suitable crystal morphology to enhance powder properties for downstream manufacturing by varying crystallization and API finishing conditions.

20.2.2.1.3 Biopharmaceutical properties

Biopharmaceutical properties of API can be characterized by two complementary well-known classification systems: the biopharmaceutics classification system (BCS) and the biopharmaceutics drug disposition classification system (BDDCS). BCS classifies drugs into four categories based on its solubility and permeability.[16] It provides a general framework to identify potential absorption barriers related to solubility and permeability and a useful guide for the rational design of immediate release (IR) dosage forms. BDDCS uses the same solubility criteria but differs in the criteria for permeability for different purposes.[17] It allows prediction of in vivo pharmacokinetic performance of drug products from measurements of permeability and solubility, including the effects of absorptive and efflux transporters and food on oral drug absorption. The importance of understanding biopharmaceutical properties of API in developing immediate-release solid products is discussed in the following sections based on drug's solubility and permeability. Considerations of biopharmaceutical properties in the design of MR dosage forms are discussed in chapter "Rational Design of Oral Modified-Release Drug Delivery Systems."

BCS class I and III compounds BCS class I compounds are expected to dissolve quickly in the GI fluid and readily cross the intestinal wall through the passive diffusion process due to high solubility and permeability. To achieve rapid and complete oral absorption, a conventional IR dosage form should be fit for purpose. Formulation and process development aspects of IR tablet and capsule formulations can be found in a series of reference books edited by Lieberman et al.[18] Although achieving fast dissolution of BCS class I drugs is usually less challenging, chemical stability, and/or manufacturing problems discussed in previous sections remain possible depending on the physicochemical and mechanical properties of the API involved and the required drug loading. Similarly, most of the potential challenges in developing IR oral dosage forms for BCS class III compounds are less likely to be related to dissolution due to its high solubility. Attempts to improve bioavailability through altering intestinal permeability using excipients at the amounts used in orally administered

formulations are ineffective due to the dynamic environment of the GI tract involving a very large surface area and effect of dilution or they are impractical because of safety concerns. Hence, the most effective way to improve oral absorption for BCS class III compounds is to increase membrane permeability through chemical modification of the API (eg, prodrugs) which is outside the scope of this chapter. Therefore, the objective of IR formulation design for this class of API is typical to make the drug available rapidly such that absorption is determined by the inherent biological properties of the API. It should be noted that certain excipients may negatively affect absorption of soluble drugs that exhibit site-dependent absorption characteristics due to their effects on GI motility and/or the transit time of drugs in the gut. For example, sorbitol has been shown to decrease ranitidine absorption. It was attributed to an increased GI fluid volume from the osmotic load of sorbitol, resulting in enhanced GI motility and reduced ranitidine transit time in the small intestine which is a primary absorption site for the drug.[19] Similar effects were observed with mannitol (an isomer of sorbitol) and polyethylene glycol (PEG) 400, All of these excipients are osmotically active at the amounts relevant to those used in pharmaceutical formulations.[20]

BCS class II and IV compounds The most common challenges posed by BCS class II and IV compounds are related to solubility or dissolution rate limited oral absorption when developing IR products. To overcome the dissolution problems, various approaches have been utilized and shown to be effective, including use of a more soluble salt, particle size reduction, improved wetting, ASD, complexation, and use of precipitation inhibitor. Forming a soluble salt is most commonly used since solubility, and dissolution rate of a salt can be many folds higher than that of the free acid or free base. Before selecting an appropriate salt for product development, solubility, dissolution rate, and other properties, such as physical and chemical stability, hygroscopicity, crystallinity, polymorphism, ease of synthesis, and toxicity, need to be thoroughly evaluated. However, it should be pointed out that use of a soluble salt for development has its limitations because it is not feasible for neutral compounds and it may be difficult to form stable salts of very weak bases or acids. For certain compounds, conversion from a salt to the less soluble free acid or free base can occur both in vitro and in vivo with pH changes. For example, a hydrochloride salt of a free base that rapidly dissolves at the low pH in the stomach may reach supersaturation on its total equilibrium solubility of ionized and unionized species. Upon emptying from the stomach, the free base may precipitate due to pH change, resulting in decreased bioavailability. In such cases, precipitation inhibitors may be used to prevent conversion from a salt to its free base, thereby, significantly improving the bioavailability of rapidly dissolving salts. Fundamental principles of this phenomenon and formulation approach to prevent precipitation can be found in chapter "Oral Formulations for Preclinical Studies: Principle, Design, and Development Considerations." Particle size reduction is also a common approach for enhancing the dissolution rate of poorly soluble drugs via increasing surface area. In most cases, API powders are micronized using cutter, hammer, roller, and fluid energy mills, which can reduce the particle size down to 1−100 μm.[21] Smaller particle size in the range of 200 nm to 1 μm may be achieved using a wet ball mill, with the milling chamber filled with grinding media and drug suspension. When particle sizes are decreased to nanometer size, the solubility may also be increased due to the Kelvin effect.[22] Therefore, use of nanoparticles can lead to an increase in both the dissolution rate and solubility for poorly soluble drugs. However, there are certain disadvantages associated with milling. Heat and mechanical impact generated during the milling process may cause physical and/or chemical instability. In addition, hydrophobic particles tend to agglomerate during size reduction, resulting in decreased effective surface area that can negate the benefit of size reduction. Excipients such as surfactants, sugars, polymers, or other excipients are often added to a formulation during processing to minimize agglomeration and improve wetting. Similar to salts, increasing dissolution in the stomach using a nanoparticle approach needs to be carefully considered to avoid the supersaturated state produced being susceptible to small changes in pH, which can cause rapid precipitation of the API. This can lead to unexpectedly low exposure with high inter- and intrasubject variability.

ASD have been successfully used to enhance oral bioavailability in an increasing number of development and commercial products of BCS Class II and IV drugs, such as, griseofulvin-PEG-dispersion marketed by Wander as Gris-PEG, Nabilone-PVP dispersion by Lilly as Cesamet, Ritonavir-copovidone dispersion by AbbVie as Norvir tablets, Vemurafanib-HPMCAS coprecipitation by Roche as Zelboraf tablets, Ivacaftor-HPMCAS dispersion by Vertex as Kalydecor tablets. Complexation using complexing agents (eg, cyclodextrin, CD) can also be used to solubilize drugs. The three-dimensional (3D) structure of the CD provides a hydrophobic cavity, which can take a drug molecule into the cavity to form an inclusion complex. The hydrophilic exterior of the CD allows the complex to interact favorably with aqueous environment. In order for the drug to be absorbed into systemic circulation, it

has to be dissociated from the drug–CD complex first. The kinetics of inclusion complex formation and dissociation between a CD and drug molecule is very rapid.[23] Dilution and competitive displacement are two major mechanisms that contribute to complex formation and dissociation process. CDs can be used to solubilize, stabilize, taste mask, and ameliorate the irritancy and toxicity of drug molecules in solid oral dosage forms.[24–26] Disadvantages associated with cyclodextrin include the potential safety concerns, high cost, and inability to accommodate high dose strength. Lipid-based drug delivery systems such as self-emulsifying drug delivery systems (SEDDS) that utilize the solubilization power of lipids and surfactants have also been used to improve oral bioavailability. It is evident that most of the formulation technologies for BCS class II and IV compounds aim to maximize the driving force for passive diffusion across the intestinal membrane through increasing drug concentration at the sites of absorption. In addition, enhanced oral absorption may also be partly a result of improved apparent permeability by saturating efflux transporters due to higher local drug concentration because most BCS Class IV drugs are substrates of efflux transporters.[17]

It should also be noted that precipitation in vivo may occur in certain cases if the drug concentration becomes supersaturated, leading to slower absorption.[27,28] The rate of drug precipitation/crystallization generally increases with supersaturation ratio (ie, the ratio of drug concentration to its solubility) and decreases with viscosity of the crystallization medium. Certain polymers such as hydroxypropyl methylcellulose (HPMC) and polyvinylporrolidone (PVP) are known to prolong the supersaturation of certain compounds. For example, PVP has been used to inhibit crystallization of sulfathiazole,[29] phenytoin,[30] oestratiol,[31] and hydrocortisone acetate.[32,33] Cellulose polymers such as HPMC have been used to stabilize the super-saturation of nifedipine,[34] albendazole,[35] hydrocortisone acetate,[36] and ibuprofen.[37] HPMC was also demonstrated to inhibit the precipitation of PNU-91325 from the supersaturable cosolvent (PEG 400 solution containing 20 mg/g HPMC) both in vitro and in vivo.[38] This had led to a fivefold increase in oral bioavailability of PNU-91325 compared to the neat PEG 400 formulation. It had been suggested that mechanisms of the crystallization inhibition might involve increased viscosity of crystallization medium, creating a steric barrier to drug molecules, adsorption to the crystals by the polymers, and/or through specific intermolecular interactions such as hydrogen bonding.[39–41] Recent studies have demonstrated that Van der Waals and hydrophobic interactions between the polymers and API play a major role in crystallization inhibition, which is discussed in more detail in chapter "Oral Formulations for Preclinical Studies: Principle, Design, and Development Considerations."

In summary, each of these approaches for improving solubility and dissolution has its pros and cons. For instance, particle agglomeration during or post size reduction can decrease effective surface area available for dissolution as discussed earlier. Physical instability is usually a major concern for ASD systems while achieving chemical stability in lipid systems can often be challenging. Complexation is incapable of providing supersaturation for increased driving force. Furthermore, in addition to solubility and permeability, many other factors are also known to influence oral absorption (see chapters: "Oral Absorption Basics: Pathways and Physicochemical and Biological Factors Affecting Absorption" and "Oral Drug Absorption: Evaluation and Prediction"). For example, chemical or enzymatic degradation in the lumen can reduce the bioavailability of certain compounds such as acid-labile compounds or proteins and peptides. Presystemic metabolism and active secretion by efflux mechanism can alter oral bioavailability, whereas uptake transporter-mediated transport may play a role in the absorption of poorly soluble or poorly permeable drugs. Many of these challenges are drug- and product-specific and thus should be evaluated based on the integrated understanding of the biopharmaceutical and physicochemical properties and required doses of individual drugs. No single formulation technology can be a solution for all.

20.2.2.2 *Consideration of drug properties in developing a "fit-for-purpose" formulation*

The approach of "fit-for-purpose" formulation development often differs for drug products of new chemical entity (NCE), generic and over-the-counter drugs. This section will focus on how one may integrate API properties into developing a fit-for-purpose formulation to support development of oral solid dosage forms containing NCEs. Development of NCEs is known to be complex, costly and time-consuming with low success rate. Many drug candidates fail to reach market due to safety, efficacy, and business reasons. In addition, the pharmaceutical industry is under increasing pressure to increase its efficiency and productivity. Therefore, it is important to design a fit-for-purpose formulation which fulfills its intended use in supporting various phases of clinical trials without requiring extensive development effort and resources. Table 20.2 summarizes objectives of clinical trials, formulation requirement, and material property consideration for developing a "fit-for-purpose" formulation.

TABLE 20.2 Development of a Fit-for-Purpose Dosage Form

Clinical trial	Objectives and characteristics	Formulation requirement	Material property consideration
Phase I	*Objectives:* To evaluate safety and tolerability. Identify a maximum tolerated dose *Characteristics:* Wide dose range Short duration (typically less than 1 month) Small number of subjects (20−100) Success rate: about 44%	Simple formulations offering a wide range of flexibility and adequate exposure to enable identification of maximum tolerated dose/establishment of sufficient safety coverage. When simple formulations cannot achieve adequate exposure, enabling technologies, such as amorphous solid dispersion or lipid-based formulations should be used to meet the primary objective of Phase I study Typically not requiring extensive stability or large scale manufacturability Speed to clinic with minimal development effort	API for reconstitution or API-in-capsule: Characterize pH-solubility profile, pKa, dissolution rate, precipitation tendency, target particle size (for dissolution rate purpose), solid form (if needed). API stability may be sufficient for shelf-life determination Tablet/capsule formulation: Additional considerations include particle size distribution needed for meeting content uniformity, solid form, minimum drug product stability needed to cover manufacturing, distribution, and duration of clinical trials
Phase II	*Objectives:* To evaluate safety and efficacy; Phase IIA: Proof of concept or mechanism Phase IIB: Identify optimal dose range for treatment *Characteristics:* 100−300 patients Duration: several months to 1−2 years Success rate: about 20%	Phase IIA: Formulations similar to Phase I or commercially viable dosage form Phase IIB: Commercially viable dosage form Sufficient stability and manufacturability required	Identify key challenges presented by API properties; Design formulation and process to overcome such challenges Understand impact of API, excipients, processing conditions, and their interactions on product quality attributes
Phase III clinical trials	*Objectives:* Pivotal studies to definitively assess efficacy and long-term safety of drug treatment compare to the current gold standard *Characteristics:* 300−3000 or more patients Duration: Several years Success rate: about 65%	Commercial dosage form and manufacturing process highly desired Stability with $\geq$2-year shelf life required Scalable and robust manufacturing process	Identify critical material attributes of API and excipients and establish control strategy. Justify formulation design based on knowledge of physicochemical, biopharmaceutical, powder properties of API and intended functionality of excipients. Identify critical process parameters and establish control strategy Understand interaction between formulation and processing conditions

Phase I clinical trials typically include a single rising dose (SRD) study, followed by a multiple rising dose (MRD) study in a small group of healthy volunteers. These trials usually encompass a very wide dose range, starting at a subtherapeutic dose, and gradually rising to a maximally tolerated dose. Therefore, formulations used in Phase I clinical trials not only need to have flexibility to cover a wide dose range but also need to achieve adequate exposure levels suitable for establishing a maximum tolerated dose. Phase I clinical trials usually are short in duration (typically $<$1 month) and involve a small batch size of clinical supply. Therefore, extensive stability data and scalable manufacturing process are not important for developing Phase I clinical trial formulations. In addition, many new drug development programs are terminated following Phase I clinical trials due to safety issues.[42] Hence, when feasible (eg, BCS class I and III compounds), a simple formulation such as API in bottle, powder in a bottle and API in a capsule can be used to enable speed to Phase I studies while conserving API and related development resources. One of the main challenges facing formulation development for Phase I studies is to achieve sufficiently high in vivo exposure, especially for drugs with limited solubility, since the higher dose in an SRD study can be many fold higher than the estimated therapeutic dose. Thus, it is crucial to understand how drug properties, such as solubility, dissolution rate and stability, may limit oral absorption. When simple or conventional formulation approaches are unlikely to provide sufficient exposure at significantly elevated dose (typically BCS II and IV

compounds), technologies for enhancing oral absorption discussed in previous sections are required, such as size reduction, lipid formulation, and ASDs. Selection of the formulation approach for a specific API needs to be based on consideration of physicochemical, mechanical, and biopharmaceutical properties of the API and their possible impact on bioavailability and manufacturing discussed previously.

Phase II clinical trials are typically designed to evaluate the short-term therapeutic effect of a new drug in patients and can often be divided into Phase IIA pilot studies and Phase IIB well-controlled studies. Phase IIA studies usually focus on proving the mechanism of action, often called "proof-of-concept" or "proof-of-mechanism" trials, whereas Phase IIB studies evaluate optimal dosage range for the treatment. Comparing to Phase I studies, Phase II trials are usually larger (including 100−300 patients) and last much longer (ranging from several months to 2 years) depending on the drug and intended indications. Therefore, formulations for a Phase II trial require a longer shelf-life and acceptable manufacturability. In most cases, Phase I formulations with adequate stability can still be used in Phase IIA studies to speed clinical trial start-up times considering the high attrition rate (about 80%) due to efficacy and/or safety issues.[42] Once proof of clinical concept is obtained, significant effort and resource are often required to develop a commercially viable dosage form for Phase IIB studies and beyond. Hence, the API properties need to be more extensively characterized for integration into designing and developing product and manufacturing process.

Phase III clinical trials are pivotal studies designed to demonstrate the potential advantages of the new therapy over other therapies already on the market. Safety and efficacy of the new therapy are studied over a longer period of time (eg, up to several years) and in more patients enrolled with less restrictive eligibility criteria (300−3000 or more depending upon the disease/medical condition studied). Phase III trial formulations and the corresponding manufacturing process need to be the same as those intended for commercial products because any formulation or manufacturing changes during or after Phase III clinical trials require justification that the product performance is unchanged, supported by appropriate in vitro testing, stability, and bioequivalence studies. In most cases, development of Phase III clinical trial formulations is a continuation of the abovementioned Phase IIB formulation. Systematic studies are usually carried out to enhance product and process understanding through identification and evaluation of critical material attributes (CMAs) of the key components including API and critical process parameters (CPPs) that are important to the development of a product and its manufacturing process. To ensure drug product quality can be met consistently during commercial production, a greater understanding of the possible impact of material and process variables and their interactions is needed to help define a control strategy of raw materials, formulation, manufacturing process, and container closure systems. The information and knowledge gained from these development studies and manufacturing experience provide scientific understanding to support the establishment of the design space, specifications, and manufacturing controls.

20.2.2.3 Product quality and performance considerations

The most important CQAs of oral solid dosage forms include drug release and stability. The former is the essential step rendering drug molecules available for absorption that would determine product safety and efficacy while the latter directly indicates whether a drug product can maintain its quality throughout shelf-life. Thus, understanding of the effects of formulation and process and their interplay on drug release and stability is the key to designing and developing a safe, effective, and stable dosage form. In developing both IR and MR oral products, formulation, and process design should be focused on (1) drug properties that influence drug release, stability, and/or processing behaviors (eg, chemical compatibility, flow, uniformity, and compaction) depending on the amount of the drug substance in each dosage unit (drug loading) and (2) potential impact of a particular process train on the physical and chemical characteristics of the drug substance that may affect product quality and performance. Specific delivery technology, formulation components, and product manufacturing processes are usually selected once a particular mode of drug release (eg, IR or MR) and type of dosage form (eg, tablet, capsule, or multiparticulates) are defined.

IR dosage forms are formulated to make drug molecules rapidly available for absorption following oral administration while MR counterparts are typically designed by altering drug-release characteristics in terms of time course and/or location to accomplish therapeutic or convenience objectives. Oral MR dosage forms include both delayed release (DR) and extended release (ER) drug products. For IR dosage forms containing soluble drugs, there are often multiple conventional formulation and process options available to meet the drug release need. The selection of the excipients and process can be made based on functional and practical considerations. However, use of excipients and process without adequate understanding can sometimes lead to issues associated with drug release

or manufacture. For example, tablet hardening on storage can lead to a change in disintegration and dissolution as a result of solid phase transition of the API or certain excipients induced by a particular processing operation (eg, milling or wetting or melting) as discussed in chapter "Crystalline and Amorphous Solids." In designing products containing insoluble drugs requiring enabling delivery technology, such as amorphous dispersions (addressed in chapter: "Rational Design for Amorphous Solid Dispersions"), an integrated formulation and process design is crucial because consistent solubility enhancement and physical stability is highly dependent not only on the API property and loading, but also on a proper combination of functional excipients and specific process technology. DR dosage forms are designed to release a drug at a later time during transit down the GI tract other than immediately following oral administration. The need for a DR delivery usually varies with the physicochemical and/or biopharmaceutical properties of drug molecules. For instance, certain drug molecules that are acid-labile require a delay of drug release in the acidic stomach to protect the drug until it reaches the small intestine. Some drugs intended for treating local conditions in the lower GI tract (eg, colitis) are also designed as a DR dosage form that requires more significant delay of drug release. In most cases, DR dosage forms are manufactured by coating single unit or multiparticulates using various enteric polymers which provide pH-dependent solubility behavior and control drug release in various parts of the GI tract. Depending on drug properties and the design objective (e.g., the extent of release delay), different polymers or amounts may be used. With regard to coating dosage forms, a pan coater is often chosen for single-unit tablets while Wurster coating in a fluidized bed is used for multiparticulates dosage forms. For ER products, different types of delivery technologies have been widely utilized over the years, such as reservoir systems, osmotic pumps, and hydrophilic matrix tablets. Each technology has its unique features, advantages and disadvantages. Drug properties and required dose usually play an important role in selecting appropriate ER technologies, their formulation composition and corresponding manufacturing process. In addition, assessing technical feasibility and designing an ER dosage form should be based on an understanding of permeability and solubility of the drug substance. These properties are particularly relevant for certain drugs for which the rate and extent of absorption differ in various regions of the GI tract, especially if absorption in the colon is relatively limited and slow. This can compromise delivery, as can reduced solubility of the drug substance further down the colon where fluid levels are low, and surface area is significantly decreased. More discussion related to rational design of ER dosage forms are available in detail in chapter "Rational Design of Oral Modified-Release Drug Delivery Systems."

Chemical stability is a critical product quality attribute that is directly linked to potency, purity, efficacy, and safety of a drug product throughout its shelf-life. Data required for assessing the stability of IR and MR solid products are usually similar to the design principles and manufacturing processes of both types of dosage forms are fundamentally the same. For most drug substances, the stability of the drug product depends not only on properties of the drug but also on formulation and process design. The first step in product design should be to understand the intrinsic stability of the drug molecule and its degradation mechanisms. Common drug degradation pathways include thermolytic (eg, hydrolysis), oxidative, and photolytic degradation.[8] Among them, photolytic degradation can usually be mitigated by reducing or eliminating light exposure using a coating or appropriate container closure systems. The most common drug degradation pathway is hydrolysis, followed by oxidation.[8] Fundamental aspects of drug degradation can be found in chapter "Drug Stability and Degradation Studies." Drug degradation is more complex in a formulated product, a multicomponent heterogeneous system. Some commonly used excipients are known to interact directly with certain drug molecules.[8] A well-known example is the Maillard reaction between a primary or sometimes secondary amine and a reducing sugar (eg, lactose/glucose). However, many interactions between drug molecules and excipients are nonspecific and not well understood, which may sometimes lead to significant stability problems in the formulation while the drug substance alone is stable. In such cases, the extent of degradation is often linked to drug loading rather than the absolute amount of drug. More significant degradation is usually associated with lower drug loading in the formulation due to a higher excipient-to-drug ratio. Drug degradation may also be enhanced due to the change of microenvironment pH created by the excipients present in the formulation. In addition, excipients can be a source of various impurities such as peroxide and metal ions which have been linked to increased drug oxidation. It is known that peroxides exist in a number of excipients such as povidone, crospovidone, PEG, polysorbates, and other modified excipients containing polyoxyethylene moieties.[43] To reduce or minimize the impact of excipients on product quality, it is necessary to assess excipient quality attributes and variability across different suppliers and different lots of the same vendor. On the other hand, certain excipient properties can also be utilized to improve product stability based on

the understanding of API's degradation pathways. For instance, bupropion hydrochloride, an antidepressant, was found to undergo extensive degradation in an alkaline environment based on a preformulation investigation. To develop an acceptable ER dosage form, weak acids or salts of strong acids were incorporated in the formulation (eg, ascorbic acid, malic acid, sodium metabisulfate, L-cystine hydrochloride, glycine hydrochloride).[44,45] These stabilizers serve to provide an acidic environment surrounding the active drug that minimizes its degradation. More recently, Oberegger et al. discovered a more stable bupropion composition for ER dosage forms through salt selection experiments.[46] However, use of highly acidic excipients may lead to degradation of the rate-controlling polymer, compromising product stability as discussed in Chapter 19, Rational Design of Oral Modified-Release Drug Delivery Systems. Thus, rational selection of formulation components based on the understanding of properties of drug and excipient as well as and their interaction is imperative.

Physical stability refers to changes in the physical characteristics of the formulation ingredients of a drug product such as phase transformation of the drug or excipients. Typical phase transformations include polymorphic conversion, hydration/dehydration, crystallization/amorphization, salt-to-free form conversion, and other mechanisms. Because the physical, chemical, and mechanical properties may differ significantly among various solid forms, these phase transformations can potentially alter the quality attributes and processing behaviors of a drug product. For example, partial or complete phase transformation may occur during processing operations[12] resulting in increased levels of crystal defects or formation of amorphous phase, which can be detrimental to chemical stability. Therefore, design of a manufacturing process should be based on the understanding of properties of drug substance and its interplay with each unit operation.

In practice, the degradation susceptibility of a drug molecule can often be predicted to a large extent from the knowledge of the drug, its molecular structure, past experience, and theoretical calculations. Thus, the formulation and process should be designed to resolve or control chemical or physical instability problems through understanding of the drug substance. For example, to minimize hydrolysis, control of moisture and microenvironment pH in the formulation, incorporation of desiccants in the container closure system, or selection of a dry manufacturing process may be considered during the product design. To mitigate oxidative degradation, an antioxidant and oxygen scavenger may be incorporated into a formulation or container closure system. To avoid or control process-induced phase transition that may affect product quality, proper selection of the unit operation or ingredient(s) that inhibit phase changes may be considered. Ideally, the solid form and appropriate excipients should be selected to reduce or minimize degradation such that no additional stability control strategies are needed in the formulation, processing, and packaging.

20.2.2.4 *Manufacturing considerations*

To ensure safe and effective products that are produced at a high level of quality, rational product and process design and commercial manufacturing all play an essential role. Despite significantly improved product quality and advances in the product development as a result of the quality-by-design initiatives over the past decade, quality issues in the pharmaceutical industry are unfortunate occurrences that have resulted in recalls, withdrawals, unprecedented drug shortages, or even risk to patients. Many of these problems have been related to manufacturing quality issues. In a recent report, the Government Accountability Office (GAO) found that approximately 40% of drug shortages resulted from quality concerns, with shortages continuing to rise in recent years.[47] Therefore, there is a clear need for enhanced understanding of the impact of manufacturing on product quality as well as the influence of product and process design for commercial production.

The production related quality or manufacturing problems encountered in the industry can sometimes be attributed to disconnects between the development and manufacturing organizations where consideration for the requirements of the next group is given a lower priority in an "over the wall" paradigm. For instance, a product development function may declare a product design "finished" once it meets the predetermined requirements for this part of an organization. Typically, these requirements focus on product quality attributes and delivery performance while neglecting other considerations, such as manufacturability, process reliability, and robustness, which can lead to inefficiency and quality issues in commercial production due to inadequate material and/or process controls or improper product and/or process design. In general, a problem encountered at the design or early development stage that is not understood and addressed will often cascade and become magnified through the subsequent development and production phases, resulting in increases in cost, timeline, and even degradations in product performance and quality. Hence, effective product development should go beyond the traditional steps of acquiring and implementing product and process design technology. It needs to consider and design patient requirements into the product while ensuring that both the manufacturing plant and the suppliers have the capability to effectively produce the product.

For a predefined QTPP, a product can be designed in different ways. A rational product development program should aim to optimize the product and process design with the production system by taking into consideration manufacturability, reliability, and robustness, manufacturing equipment, processes, capability, capacity, experience and expertise, material handling systems and supplies, among others.

The key to the success of such an integrated program hinges on the ability to obtain knowledge for making sound decisions to enable effective transition to the next phase. To avoid problems with serious ramifications later during scale-up or after product approval, it is important to adopt a knowledge-based approach by ensuring the right level of knowledge at each key decision point. To ensure that manufacturing considerations are taken into account earlier, when costs to anticipate and identify problems and make product and/or process design changes are significantly lower, it is crucial to start with the end in mind and to use concurrent scientific and engineering principles to capture design and manufacturing knowledge and conduct risk assessment with a proactive rather than a reactive focus. In particular, a key measure of manufacturability is whether the critical manufacturing processes and key properties of materials are in control and product reliability is achievable. This knowledge can facilitate informed decisions to develop a product by taking steps to ensure the product and process design is stabilized early so product performance and manufacturability can be demonstrated before production. For example, achieving consistent dissolution performance of a coated ER dosage form during scale-up or upon storage is often challenging. If product quality standard and process control are defined in the development phase without adequate understanding of formulation behavior, critical properties and variability of the rate-controlling polymers and critical process controls, there is a risk that products being transferred to production could only be made to work at the pilot scale by highly skilled technicians. But they could not be manufactured to meet quality specifications at commercial scale, or might need to rely on a "use test" for individual batches to pass dissolution specifications. In designing a tablet product made using direct compression process, basic studies of segregation tendency and compaction properties of various formulation blends are often carried out in the laboratory for the risk assessment associated with content uniformity and/or compression. However, it is more important to start formulation and process design by first understanding inherent properties of the API and available excipients, their interactions with one another and with the chosen unit operations and the anticipated effects of scale change. It is equally important to assess natural variations of critical material and process variables, how they may impact quality, manufacturability and the ability to control at the production scale. This type of knowledge not only can help guide formulation design, but also facilitate a sound decision to change the manufacturing process if the selected formulation and process design is found to be too sensitive to tolerate expected variations of critical variables of key components, process parameters, and/or equipment in the commercial production.

In the current patient-focused quality, manufacturing and regulatory environment, integrated product and process design for both functionality and manufacturability based on product and process understanding is critical to the development of high quality, highly functional products that can be effectively manufactured. A sound product and process design should balance product quality with design effort and ability to compensate for uncertainty in the raw materials, manufacturing, and testing. Process robustness and reliability activities should start at the earliest stages of product and process design and be integrated into the overall product development cycle. It should also continue throughout the life of the product. Although it may sometimes require additional effort early in the development process, the integration of product and process design through science and technology and improved business practices will result in a more reliable and manufacturable product to better meet patient needs, a quicker and smoother transition to manufacturing, a higher production efficiency and a lower total product life cycle cost.

20.2.2.5 Selection of dosage form and production method

Dosage form development involves design and development of a product with a predefined QTPP and manufacturing process. The most common solid dosage forms are single or multiunit tablets or capsules manufactured using various established processes and unit operations. The general aspects of these dosage forms, related technologies, process development, and in-process control (IPC) have been extensively addressed[18,48] and also discussed in details in the relevant chapters of this book. In general, processes for tablet manufacture include direct compression, wet granulation (high and low shear or twin-screw extrusion), dry granulation (eg, roller compaction, slugging), and melt granulation or thermoplastic pelletizing (eg, high shear, fluid bed, melt-extrusion, spray-congealing). Processes for producing round solid beads encompass extrusion-spheronization, drug layering of nonpareil seeds, spray-granulation, powder layering, spray-drying, and spray-congealing. Conventional compression processes are also utilized to prepare mini-tablets using either small conventional tooling or

multihead tooling (eg, ≤ 3 mm) depending on the tablet size and throughput considerations. Coating of tablets or beads is most often carried out in a pan coater or fluid-bed technology depending on the nature of the coating substrate (eg, size, tensile strength). Coating by compression (tablet-in-tablet) requires special compression machines. When a more sophisticated process is required for products of increased complexity, for example, multilayered tablets, compression coating, mixed beads, and/or mini-tablets, emphasis should be placed on the increased process and product understanding throughout development lifecycle to ensure successful development from small to commercial scale.

In the product design phase, often there is more than one delivery technology, formulation approach, or production method that is suitable for developing a viable product with predetermined quality and performance for a specific API. Rational selection of a dosage form and associated drug delivery technology, formulation, and manufacturing process needs to be based on a concurrent and integrated evaluation of a multitude of factors discussed in the previous sections, such as API characteristics, material properties, processing operations and their interaction, variations, process robustness, efficiency, and cost. Also, the use of conventional manufacturing processes and equipment is highly preferred whenever feasible.

20.2.3 Product and process understanding

A sound dosage form design provides manufacturers with a quality product and helps to ensure manufacturability and optimum productivity. Ideally, physicochemical and mechanical transformation processes of the raw materials that ultimately result in the desired product should be based on a mechanistic understanding of the materials and processing operations. However, the diversity and complexity of pharmaceutical ingredients, combined with the unit operation paradigm utilized in pharmaceutical manufacturing, usually make a thorough understanding difficult. In many cases, knowledge gained appears to be material- and equipment-specific (dependent on machine design, geometry, model, manufacturer, etc.), making it challenging to develop universal models for direct application to different scales, processes, and products. Nevertheless, it is both important and possible to overcome these obstacles and develop consistent, high-quality product and robust manufacturing process through the understanding of the API, formulation, process, sources of variability, and risks to product quality. The most important element in achieving a high level of product and process understanding is to utilize the scientific knowledge discussed in the previous sections and apply a systematic approach through a combination of design of experiments (DOEs), studies that lead to scientific understanding, engineering principles, empirical, semi-empirical, and mechanistic models where possible. Specifically, in the product development process, selection of an appropriate delivery technology, composition, and corresponding process option should be based on a comprehensive analysis of information of (1) target product characteristics (required dose, release mode, type of dosage form, etc.); (2) physicochemical, mechanical, and biopharmaceutical properties of the API; (3) formulation technology and level of complexity; (4) processing technology and associated operations, equipment, and sophistication of production; (5) raw material attributes, inherent variability, and potential interactions (eg, drug-excipient, excipient-excipient); and (6) interrelation between materials and unit operations. This type of information integration can be used to anticipate and evaluate the possible impact of formulation on processing or vice versa, including how product characteristics may dictate process choice or how processing operations may limit formulation selection. It is also required for assessing influence of formulation/processing on in vitro and in vivo performance. Examples have been provided in the previous sections to illustrate the importance of understanding API properties to achieve product performance and ensure quality. Many other examples of adopting a holistic approach to the development process can be found throughout this book. For instance, in manufacturing tablets, components present in high percentages (eg, drug or polymer) in the formulation often merit an increased level of scrutiny and quality control because they often influence or even dominate the processing behavior of an individual unit operation, such as fluidization, granulation, blending, or compaction. Assessment of compressibility of a tablet formulation requires characterization of the deformation and bonding properties of the major components and relationships with compression parameters (force, turret speed/dwell time, precompression), and moisture content for predominantly plastically deformed materials. In drug layering processes, low process efficiency resulting from drug–polymer interactions can be resolved more effectively by replacing the interacting excipient when compared with the attempt to increase efficiency via optimizing processing conditions.

In developing and manufacturing MR dosage forms, understanding and controlling critical properties and the natural lot-to-lot variability of rate-controlling materials within or outside the supplier's specifications are often necessary to minimize their potential impact on processing, drug release and/or in vivo performance. In recent years, the potential implication of solid phase change during processing is another area

that received increased attention.[7] For example, solid phase transition from the stable crystalline to the less stable amorphous drug caused by processing conditions can directly lead to chemical instability of the product.[49] Tablet hardening with age may adversely affect the disintegration and dissolution rate during storage.[50] Depending on the material properties and processing conditions, the change of tablet hardness may be attributed to recrystallization of a physically stable phase from a small portion of the metastable or amorphous form of either the drug substance or excipient (eg, mannitol) generated during processing. It should be noted that this type of phase transition is often a function of material characteristics, composition, type and processing condition of a specific unit operation. Its impact on dissolution may also depend on product design and associated dissolution mechanism. Through characterization of the materials and their interaction with the dosage form and process, the solid phase changes can be anticipated and avoided if they are undesirable. They may also be controlled and utilized to enhance product performance (eg, dissolution).[7]

In the process of applying a systematic approach to the concurrent design of products and processes, one of the most commonly employed tools is DOEs in addition to the use of prior knowledge. DOE is a structured, organized method for determining the relationship between factors (Xs) affecting a process and the output of the process (Y). For pharmaceutical processes with measurable inputs and outputs, it can be used to guide the optimal selection of inputs and analysis of results in parameter design, problem solving, and robustness study. Its basic concepts are available in the Experimental Design section of a free electronic textbook: StatSoft at http://www.statsoft.com/Textbook/Experimental-Design. The fundamental principles, techniques, and applications of DOE can be found in many books written on this subject.[51] In pharmaceutical development and manufacturing, appropriate application of DOE has proven valuable in maximizing information gained with minimum resources, identifying critical factors among a multitude of variables, solving problems in a complex system consisting of many confounding factors, defining, optimizing, and controlling product and process variables.[52–56] In fact, DOE is so widely applied in today's pharmaceutical research and development (R&D) and operations that there may be a perception that running experiments using DOE is equivalent to product and processing understanding. This, in part, is a result of the historical successes of DOE across multiple industries, and partly due to a lack of adequate understanding of the intricacies and pitfalls of designing and executing such experiments in the study of complex pharmaceutical systems and their performance in the multivariate environment and/or highly variable clinical settings. For example, understanding the measurement precisions of inputs versus those of outputs and the need for repeat observations, identifying confounding factors are not trivial, while putting together a DOE design table is a very simple task as all of the well-known classes of standard designs can be generated automatically by software once the number and nature of input variables and the responses are defined. It is important to point out that an incorrectly or improperly designed DOE not only can result in wasted resources and time but often generate data that are confusing or misleading. Some of the common problems associated with DOE experiments in dosage form development or investigation include (1) choosing the inappropriate design, variables and, in particular, their ranges; (2) running uncontrolled experiments due to obliviousness of confounding factors; (3) generating unreliable or inconclusive data resulting from inadequate repeat measurement, sample size, or lack of understanding of the signal-to-noise ratio (SNR). Therefore, it is imperative to gain a clear understanding of the study factors and responses, variability, possible confounding and interacting factors, analytical variation before selecting appropriate input and responses, the range of investigation, and experimental design. More specifically, when choosing and implementing an experimental design, it is vital to first acquire a basic knowledge of the studied system and utilize multidisciplinary principles and expertise (eg, pharmaceutics, chemistry, engineering, statistics, among others) for identifying test variables and defining levels. Moreover, knowledge of the difference between control and noise factors is also critical. In particular, it is very important to carefully select a small number of experiments under controlled conditions and take into account the following interrelated steps:

1. Define an objective to the investigation, for example, identify important variables or find an optimum.
2. Determine the design variables that will be controlled and their levels or ranges of variation, as well as variables that will be fixed.
3. Determine the response variables that will be measured to describe the outcome and examine their precision.
4. Among the available standard designs, choose the one that is compatible with the objective, number of design variables and precision of measurements.

It should be emphasized that DOE design is, in most cases, a strategy to gather empirical knowledge, that is, knowledge based on the analysis of

experimental data and not on theoretical models. Thus, it is mainly useful for investigating a phenomenon, not the underlying mechanistic cause of a problem. Thus, use of experimental design strategy in conjunction with the application of the scientific and engineering principles, techniques, and practices presented throughout this book is essential to the effective and efficient development of robust product and process from laboratory to production scales.

20.2.4 Process scale-up and optimization

Following the completion of formulation and process design, essential steps required for final product commercialization include process scale-up and optimization. Process scale-up is referred to as the process transfer from small scale (laboratory scale or pilot scale) to a larger scale (pilot scale and/or production scale).[57,58] Process scale-up requires careful planning and effective selection of equipment, process parameters, and a control strategy based on product and process understanding and experience gained during development. To warrant a robust manufacturing process that is capable of consistently producing safe and effective product in the commercial setting, the influence of equipment, material and process attributes as well as variability on CQAs should be understood thoroughly.[2,3,59,60]

Scale-up is often defined as increasing batch size.[57,58] Depending on the nature of the process, increasing batch size does not necessarily imply the enlargement of processing volume of each unit operation. Examples of increasing batch size without enlargement of process volume of the unit operation include combination of multiple small granulation batches and extending the compression time using the same tablet press. In these cases, the transition to larger batch size is relatively orderly and straightforward. However, when equipment changes are required to increase batch size, process scale-up could become more complex. This is due to the fact that scaling-up of many parameters of pharmaceutical batch processes is often nonlinear and lacks a reliable scale-up model. Thus, it is difficult to predict large-scale process performance or define process parameters based on data and experience with small-scale equipment without conducting process experiments at an intermediate or the intended full scale.[2,57,58,60]

Under the QbD paradigm, process scale-up and optimization should follow a systematic, scientific, and risk-based approach.[2,3,59,60] The typical scale-up and optimization process can be illustrated in Fig. 20.2.

The first essential step for a successful scale-up should be thoroughly evaluating critical manufacturing variables for each unit operation, including CMAs and CPPs identified from small or intermediate scale experiments.[2,3,59,60] Additionally, equipment evaluation between scales to understand and identify the capability and limitation should be performed. Potential scale-dependent process parameters should also be identified during this step.

Scale-up experiments can then be designed to (1) verify the small scale results and (2) establish design space and controls strategy at the large scale. Various sources of variability should be properly included in experimental design to better understand the process robustness. Ensuring product quality during scale-up, maintaining geometric, thermodynamic, or kinetic/dynamic similarity between scales are basic principles that should be employed.[57] When the dimensional analysis can be applied, dimensionless numbers developed based on these similarities are largely independent of scale.[57,61,62] It is believed that two processes can be considered similar when they take place in similar geometrical space and all dimensionless numbers describing the process have similar numerical value.[57,61,62] According to this theory, the scale-up procedure would be: (1) expressing the unit operation into set of dimensionless numbers, and (2) matching the numeric value of those dimensionless numbers among different scales. An example of often used dimensionless number is the Froude number, which has been applied to the scaling of granulation process.[63] Computer based modeling, such as Finite Element Model (FEM), Discrete Element Method (DEM), is also starting to gain interest for scaling processes. However, due to the complexity of pharmaceutical manufacturing processes, the development of quantitative, first-principle-based dimensionless numbers and models that accurately describe the system are often unavailable. Hence, empirical/semiempirical approaches have been utilized to determine the process set points and ranges. When process parameters are scale-dependent, two types of scale-up approaches have been used: parameter-based and attribute-based. In practice, a combination of both strategies is often utilized.[64] One example of empirical scale up is to match the tip speed among intermediate bulk containers (IBC) of different sizes for dry blending/mixing process of free flow materials. Such a scaling principle provides a simple and quick estimation of the process operating range. Prior knowledge could also play a significant role and should be utilized when designing the scale-up experiments. When conducting process scale-up trials, more frequent testing than would be typical during routine manufacturing is usually needed to understand all potential variabilities and their impact. Defining proper sampling plans and acceptance criteria are both

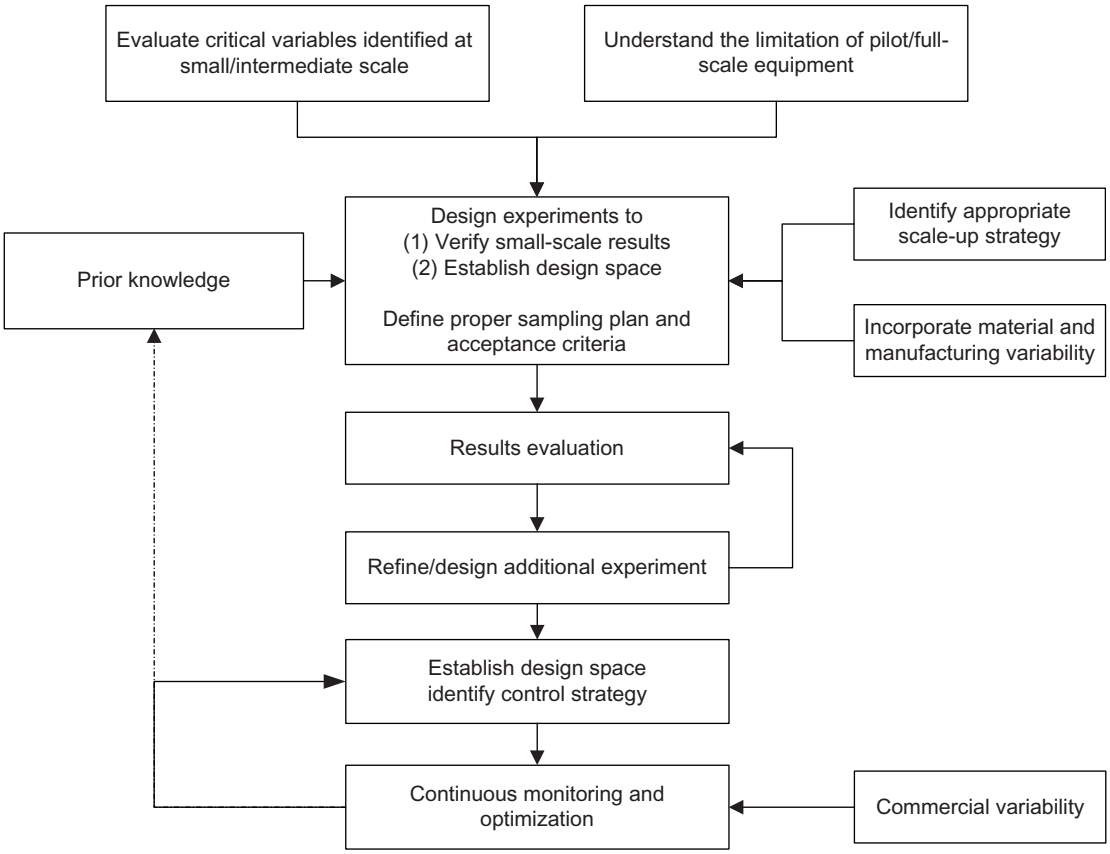


FIGURE 20.2 Typical road map of process scale-up and optimization.

critical and should be aligned with the identified CQAs and/or IPCs to ensure predetermined product quality and performance remain unchanged. On-line process measurement, if available and reliable, can provide real-time results without potential sampling bias and process disturbances.[2,65] As process scale-up is an inherent part of product process validation lifecycle, appropriate monitoring programs for quality attributes and process parameters should be established at this stage.[66,67] Those monitoring programs can serve as the basis for further process optimization.

Prior to commercial production, it is essential to build confidence in process consistency within the established process ranges with respect to performance indicators and product CQAs, However, in most cases a significant amount of data are usually unavailable at the development stage due to time, resource and material constraints. Thus, product and process understanding along with proper statistical tools are needed as a product moves through different stages of its lifecycle to gain increased understanding of manufacturability and process robustness. For example, process capability expressed as a process capability index (eg, Cpk) or process performance index (eg, Ppk), is a measurable statistical property of a process to the specification (see chapter: "Commercial Manufacturing and Product Quality" for definition). At the beginning of production, Cpks for all critical product measurements should be calculated to determine the general ability of the process to meet specifications. However, to obtain meaningful results, data should include the normal variation of production conditions, materials, and operators. In general, more than 25 subgroups of results are recommended.[68] Performing capability analysis with data or results of an inadequate number of batches could potentially lead to skewed analysis due to sampling variability and limited sample size. Throughout the lifecycle of the product, since a much wider range of variability will be encountered, process performance should be evaluated from time to time to enhance process understanding and identify optimization opportunity. After commercialization, an important part of process improvement activities is variability reduction through assessing process capability as well as material and process control. Although manufacturing efficiency, safety, and product yield are rarely considered critical to product quality attributes, they indirectly reflect a degree of process certainty and control, therefore, should be carefully

evaluated during scale-up and optimization. For example, inconsistent or lower than expected product yield indicates manufacturability issues. In such cases, process and control strategy should be reexamined and improved accordingly.

In summary, process scale-up and optimization is essential to pharmaceutical commercialization. It is an integral part of product process validation lifecycle. The experimental design for scale-up and optimization should be based on product and process understanding. An appropriately designed process scale-up and optimization will provide a high level of confidence in manufacturing robustness and consistent product quality. Detailed discussions of process scale-up can be found in individual chapters dedicated to various pharmaceutical processes in this book.

20.3 TECHNOLOGY TRANSFER

20.3.1 Technology transfer overview

In the pharmaceutical industry, "technology transfer" may broadly refer to the business processes required for successful progression from drug discovery to product development, clinical trials, and full-scale commercialization. For the development and transfer of knowledge and manufacturing technologies of finished solid dosage forms, it is usually defined as transferring the product out of the development laboratories and the process development pilot plant into full commercial-scale manufacturing facilities for new product launch. Technology transfer can also refer to transferring a commercial product from existing commercial sites to other commercial sites, also referred to as "site to site transfer." This section mainly covers the key points for the technology transfer related to new product launch. However, most of the principles described here can also be applied to site transfer.

Appropriate technology transfer includes the transfer of technical knowledge of the product and process, analytical test methods from a sending organization to a receiving organization, as well as the establishment of the related quality control system and documentation for the new product at the receiving organization. The sending organization is typically an R&D organization or a development partner that has developed the technology. The receiving organization is normally a commercial manufacturing site (for product/process transfer) or a quality control laboratory (for test method transfer). This process is important to ensure that product quality and performance developed during R&D remain unchanged in full-scale commercial product manufacturing and can be maintained during routine commercial manufacturing after launch.

Technology transfer is a complex process that typically involves multiple functional areas. The ultimate goal is to effectively transfer manufacturing process, specifications and controls, as well as analytical methods based on sound scientific and engineering principles, while striving for optimization of efficiency, flexibility, and capability of the process. Technology transfer should also be a continuous knowledge/experience transfer and information exchange between parties, rather than a one-time event. Problems often arise at interfaces when the transfer process is not properly planned, communicated, and executed, and can result in a high rate of batch rejections or even a delay in bringing new products to market. Thus, effective and efficient technology transfer has been, and will continue to be, critical to the success of the pharmaceutical industry. Technology transfer may include API process transfer, drug product process transfer, and analytical methods transfer. Analytical methods transfer is normally part of the API or drug product transfer to support the manufacturing site change. The API process transfer includes transfer of the API synthesis of intermediates and final products as well as test methods for raw materials, starting materials, reagents, catalysts, intermediate and final products. A lab run using reagents from the commercial site is normally the first step in the process transfer. A solvent run is then conducted at the commercial site using the same process train and operating conditions without the use of solids and other raw materials except solvent. It is intended to test the manufacturing procedure and equipment capabilities without consuming expensive and limited raw materials. Critical process control parameters such as temperature, solvent charge quantities, mixing and material transfer rates, can be studied using a solvent run. Following the solvent run, engineering run(s) are typically conducted to fine tune the process parameters. The experimental data together with the data generated in the R&D pilot plant and at lab scale can be used to justify the process ranges and IPC limits. After the process ranges are justified, IPC and final product specifications are established, and the API process can be qualified at the commercial sites. The International Society for Pharmaceutical Engineering (ISPE) provided a detailed guide on the requirements for API transfer.[69] This section will mainly focus on the drug product transfer.

20.3.2 Technology transfer of drug product

A technology transfer of drug product normally includes the transfer of the product/process as well as

the drug product test methods. Sometimes, the API test methods also need to be transferred to the DP manufacturing site since API release testing might need to be performed to ensure quality control of the incoming materials. Technology transfer of drug product is normally initiated when the compound reaches Phase IIb in drug development. This is often the time when planning of commercialization of the compound is initiated, and a commercial manufacturing site is selected. During initiation of the technology transfer, the overall project scope is defined, technology transfer lead, stakeholders and team members are identified, and high-level technology transfer milestones are established. A technology transfer team typically contains members from multiple functional areas, such as pharmaceutical and analytical R&D, operations (manufacturing and engineering), quality, regulatory affairs, supply chain, and program management. As with any multifunctional team, clearly defining roles and responsibilities of team members is the key to success. The technology transfer team lead plays a key role in the success of the transfers and is expected to be an individual with appropriate technical experience and knowledge from drug development and manufacturing, project management skills and experience, and excellent communication/negotiation skills. Stakeholders include the management from the key functional areas who are funding or affected by the technology transfer (eg, R&D, manufacturing site, quality, and regulatory and supply chain). The high-level technology transfer milestones normally include:

- Site readiness, especially when a new or modified facility or new equipment is introduced
- Completion of test method transfer to the receiving lab(s)
- Completion of process transfer to receiving site
- Phase III clinical supply readiness (if Phase III supply is manufactured from the commercial site)
- Completion of stability batch manufacturing at the receiving site (site specific stability batches)
- Completion of demonstration, or confirmation, batch at the receiving site
- Completion of process validation at the receiving site
- Regulatory filing submission

20.3.2.1 Technology transfer planning

After forming a technology transfer team and setting up technology transfer milestones, the technology transfer team needs to develop a detailed technology transfer strategy/plan which typically includes project scope, key deliverables and timeline, project resource requirements, transfer strategy, gap assessment, and acceptance criteria.

Project scope needs to define clearly what is in the scope of the specific transfer. For example, the transfer may or may not include transferring the final packaging operation or an alternate API source may or may not be in the scope of the transfer. Key deliverables for a technology transfer normally include timelines for the key milestones agreed upon. A full list of team members should be provided in the project resources section of the technology transfer plan.

The transfer strategy includes the master plans for the transfer of the product and process knowledge to the manufacturing site, test method transfer, stability and shipping, packaging, market and material sourcing, regulatory strategy and facility and equipment strategy. Planning of the product and process knowledge transfer includes detailed action plans and timelines to further develop the commercial process/formulation, scale up and transfer to the commercial site, manufacturing of registration, demonstration/confirmation and validation batches. API specifications, excipients specifications, and drug product specifications need to be established at the receiving site with the related knowledge transferred from the sending site. A site specific hold time study is recommended to be performed during the technology transfer to justify the process hold time of product at each stage of the process. Since the hold time study normally takes time to complete, proper planning at the early phase of transfer is recommended. Cleaning procedures and cleaning validation are typically site specific. The cleaning procedures from the sending site should be compared with the procedures at the receiving site to assess if the cleaning procedures at receiving site need to be modified to meet the requirement of the new compound. Packaging processes (including bulk packaging and finished product packaging process) should be considered as part of the process transfer as appropriate. Different packaging material might require different packaging conditions. When new packaging material is introduced, thorough packaging process development work could be helpful to establish the optimized packaging process parameters for the new material. Test method transfer strategy includes a list of methods that need to be transferred, the sending and receiving sites of each method, and transfer strategy for each method. Depending on the capability of the commercial site, some methods might need to be transferred to other third party labs. In many cases, the release lab is different from the labs that perform the stability study and testing. Therefore, test method transfer to stability labs should also be considered. The methods being transferred include API, excipients and drug product release methods, cleaning methods including testing for residual API

and detergent, as well as any required in-process testing methods. The testing sites to support formulation/process development, registration, clinical lot release and stability should also be planned and selected to ensure resources are properly allocated. Stability strategy contains the plan for primary stability, site stability, and any supportive stability studies (eg, in-use stability and bulk stability) required for a product launch. A matrix or bracketing approach may be used where appropriate. A shipping study justifies the temperature excursions during shipping and may include a simulated or real time shipping study to evaluate the physical integrity of the drug product during shipping. The packaging strategy includes the packaging configuration planned for bulk API, bulk DP, finished DP as well as the packaging site. Selection of packaging configurations needs to consider both chemical and physical stability of the drug substance and drug product. Market and material sourcing strategy includes the targeted markets for the DP and potential material vendors. Vendor qualification activities need to be planned for new vendors. Plan for alternative suppliers should also be addressed especially for critical excipients or supplier at higher risk. The regulatory strategy includes regulatory assessment of data requirements for filing in countries where the drug product will be marketed as well as planning to meet expectations during any anticipated GMP inspection or prior approval inspections. The regulatory strategy is closely linked to stability strategy since the stability requirements might be different from country to country. Some country-specific documentation may take time to prepare and should be listed in the regulatory strategy and planned accordingly. In a transfer where a new facility needs to be built, and new equipment needs to be purchased, the technology transfer planning should also include capital purchase and facility engineering planning. The environmental differences (eg, temperature and relative humidity control) from the receiving site and sending site also need to be considered to ensure product quality is not impacted during manufacturing. Capital items may include the cost of the facility, IT system investments, equipment installation and qualification expenses, lab equipment purchasing, certification, among others. Expense usually includes rearrangement of the current facility, demolition, engineering, certification, among others. The initial manufacture estimate (IME) often includes employee costs, general production support, disposal costs, process cost, among others. A gap analysis for technology transfer needs to be performed in the early phase of the transfer to capture all the identified gaps with a remediation plan. The gap assessment normally includes potential process/analytical equipment gaps, environmental control gaps, process/product knowledge gaps, etc. The final piece of the technology transfer plan is acceptance criteria. The metrics defined in Table 20.3 provide a baseline to measure technology transfer success.

TABLE 20.3 Metrics for Measuring Transfer Success

Metric	Description
Drug product facilities	DP site(s) have necessary equipment and procedures in place for manufacturing, testing, approving, cleaning, and packaging per project timeline
Analytical methods	Methods are transferred and ready for use per project timeline for API, excipients, drug product, and cleaning
Drug product development	The process operating space is defined and justified The process is properly scaled up to meet commercial-scale requirements Process validation requirements are incorporated in the product development to ensure successful process qualification and robust commercial manufacture Packaging requirements are established to ensure product stability
Regulatory strategy	The regulatory strategy is defined and incorporated into the tech transfer requirements and project timeline so the filing occurs on schedule
Launch readiness	Quality system and product-related manufacturing and testing documentation are in place and commercial supply chain is ready in time for DP approval from regulatory authority. Adequate supply of product is available for launch

The technology transfer plan is a living document and should, if necessary, be revisited and revised (with appropriate version control) periodically during execution of the transfer.

20.3.2.2 *Execution of technology transfer*

During the execution phase of the transfer, a detailed project timeline based on the technology transfer plan should be developed. Effective communication between all functional areas throughout the execution phase is crucial. Documentation during the execution phase is not only important to ensure a smooth transfer but also required by the regulators. Table 20.4 provides an example of the documents that need to be available at the receiving sites for the technology transfer.

Additional information for pharmaceutical technology transfer can be found in the guidelines published by World Health Organization in 2011[70] and by the

TABLE 20.4 Examples of Documents Needed at Receiving Sites

Category	Documents
API	API specification and any additional characterization that have potential impact on drug product
Formulation/process	Formulation development report (including formulation declaration, justifying function, type and amount of components, and API characteristics)
	Development report for Phase II DP manufacturing process including the following information:
	Process justification identifying and justifying the critical process parameters and critical quality attributes Batch summary of historical batches Equipment scale and its operating parameters and ranges Nonconformities and investigations API/excipients used and COAs Summary of batch release results and in-process testing results Special storage and product handling Special hold time requirement Yield of each step Manufacturing batch records (blank and executed) Site SOPs Cycle time of each step
Facility and equipment	Facility requirements: Containment Environmental Utilities Room temperature and humidity controls IQ/OQ documentation Equipment requirements Any modification to the commercial equipment/instruments and computer systems IQ/OQ/PQ documentation Equipment selection justification Material transfer equipment Capacity Controls Operating ranges Calibration/validation ranges PAT Cleaning procedures
Environmental health and safety	• Dust explosion data • Special emission controls • Waste streams (description, composition) • Toxicity information • Process hazards
Analytical	• Stability reports listing the current data and status of ongoing and completed stability studies • Validation or transfer documentation (eg, failure mode and effects analysis (FMEA), risk control strategy, gap assessments) necessary for method understanding • Analytical method development reports • Availability of samples for the test methods transfer • Hands on training and best practices prior to test method transfer • Analytical equipment setup • Specification and specification justification document (including drug product and excipients)
Cleaning	Cleaning method and cleaning validation report
Packaging	• Packaging selection/justification report • Type of packaging and specifications: • Primary packaging for the finished product • Bulk packaging • In-Process material packaging • Hold time studies • Vendors for packaging • Environmental controls during packaging and storage • Shipping study or simulated shipping study

ISPE in 2014, respectively.[69] These guidelines present a clear standardized process for transferring technology between two parties and recommend a minimum base of documentation in support of the transfer request. They define key terms with consistent interpretation that enables the capture and sharing of essential knowledge. These guidelines are useful from the earliest phase of a product's lifecycle to postapproval transfers, and provide guidance and insight into the essential activities and documentation required to move a product, process or method from one unit to another. The information is equally applicable to ethical and generic products, as well as technologies originating from any region around the world.

20.4 CASE STUDIES

20.4.1 Influence of material properties and processing conditions on tablet capping

Capping is one of the mechanical defects that can occur in the tableting process and it can cause catastrophic failure of the compact. Capping is defined as the separation of the top or bottom curvature part of the tablet from the tablet body either partially or completely. It can occur as the tablet ejects out of the die, or during subsequent handling or hardness testing. Tablet capping may be influenced by formulation material properties, process conditions and tooling shape. In recent work published by Akseli et al.,[71,72] the authors developed a nondestructive ultrasonic method to correlate material and process variables quantitatively to tablet capping tendency. They studied a variety of powder blends with a wide range of material properties. These blends consisted of different levels of two model drugs (acetaminophen (APAP) and ascorbic acid) with either brittle fillers (lactose monohydrate, or Tablettose 80) or plastically deforming fillers (microcrystalline cellulose, or Avicel PH 102). In addition, they also studied 10 commercial production blends for which there was extensive manufacturing knowledge of capping tendency. To compare the capping tendency between commercial products with the experimental blends made at lab scale, all blends were compacted under the same controlled conditions using Presster, a compaction emulator. The number of tablets that capped during ejection and hardness testing for each powder blend was counted and used to calculate the formulation capping tendency. A nondestructive ultrasonic tester was used to analyze the acoustic properties of pressure and shear waves in the tablets from which it was possible to derive elastic properties such as Young's modulus E (ie, describes the tendency of a material to deform along an axis when opposing forces are applied along that axis) and shear modulus G (ie, describes the tendency of a material to shear when acted upon by opposing forces). Tablets were tested along axial and radial directions and dimensionless ratio $EG^{E} = E_{axial}/E_{radial}$ and $EG^{G} = G_{axial}/G_{radial}$ were calculated and used to reflect how elastic properties vary along axial and radial directions of tablets. The key conclusion included that: when the ratios EG^{E} and EG^{G} (referred to as *EG* ratios later) approach unity the tablet is expected to be more homogeneous, and capping is less likely to occur: As the *EG* ratios deviate from unity, heterogeneities in the tablet are expected to be more pronounced, and capping is more likely to occur. Fig. 20.3 shows time of flight (TOF) for longitudinal (pressure) and transverse (shear) ultrasonic pulses in the tablet acquired for one of the test products (Product-9) in both axial and radial directions for a commercial product. The calculated *E* and *G* ratios of 0.98 and 0.95 indicate no risk of tablet capping with this commercial product.

Based on the study, a very good correlation between the *EG* ratios and tablet capping tendency could be obtained among 26 lab and commercial samples with widely ranging material properties. As shown in Fig. 20.4, no capping was observed either during ejection or hardness testing when EG ratios are greater than 0.65, which was defined as the green "acceptable region." When the *EG* ratios are between 0.45 and 0.65, capping was occasionally observed during hardness testing, which was in the yellow "moderate region." When the *EG* ratios are less than 0.45, tablet has high risk for capping, which was shown in the pink "poor" region.

Material properties usually have the most significant influence on the capping tendency of formulation blends. Take APAP blends, for example, it is well known in the literature that APAP containing formulations are prone to capping due to poor compression characteristics of APAP.[73] From Fig. 20.4 it can be seen that when APAP was blended with Tablettose 80, a commonly used filler which undergoes brittle fracture during compression, the resulting powder blend exhibited high tendency for capping when the APAP loading is as low as 5%. However, when APAP was blended with microcrystalline cellulose, another commonly used filler which undergoes plastic deformation during compression, the resultant blend did not show a tendency to cap even at a drug load of 20%. When a different model drug, ascorbic acid, was tested, binary blends with Tablettose 80 containing 1–10% API did not exhibit capping tendency. However, when drug loading was increased to 20%, high capping tendency was observed. Similar to the APAP example, use of plastically deforming filler

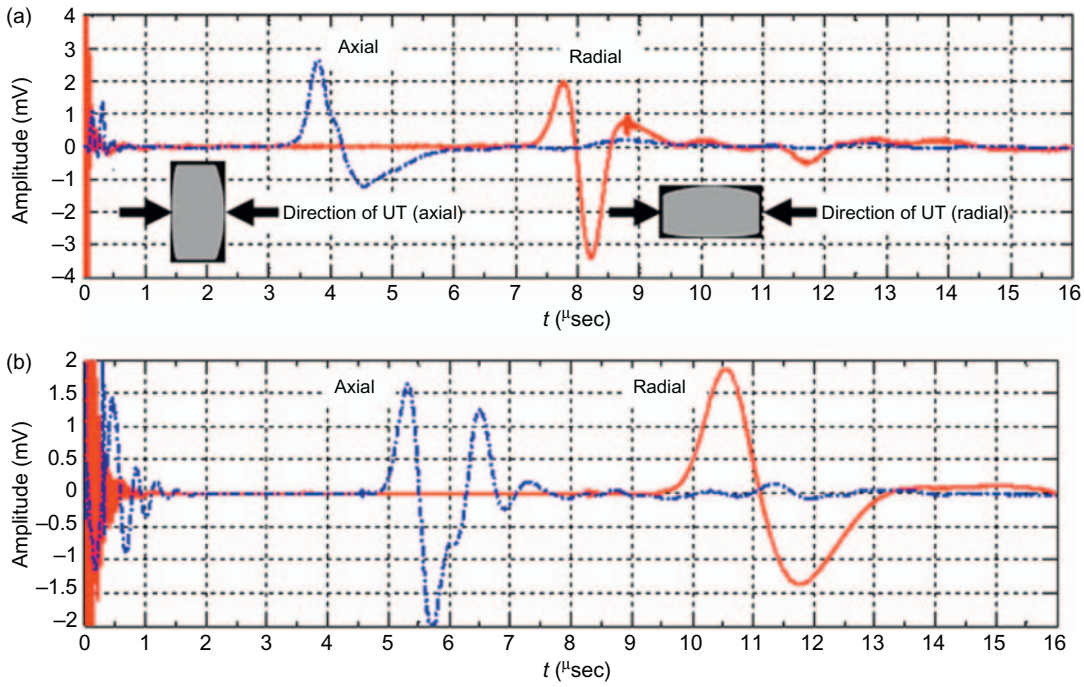


FIGURE 20.3 Comparison of the axial (blue dashed line) and radial (red solid line) TOF of longitudinal (pressure) (a) and transverse (shear) (b) ultrasonic pulses in the tablet acquired for Product-9. The capping rating of this product is 3 (acceptable).[71]

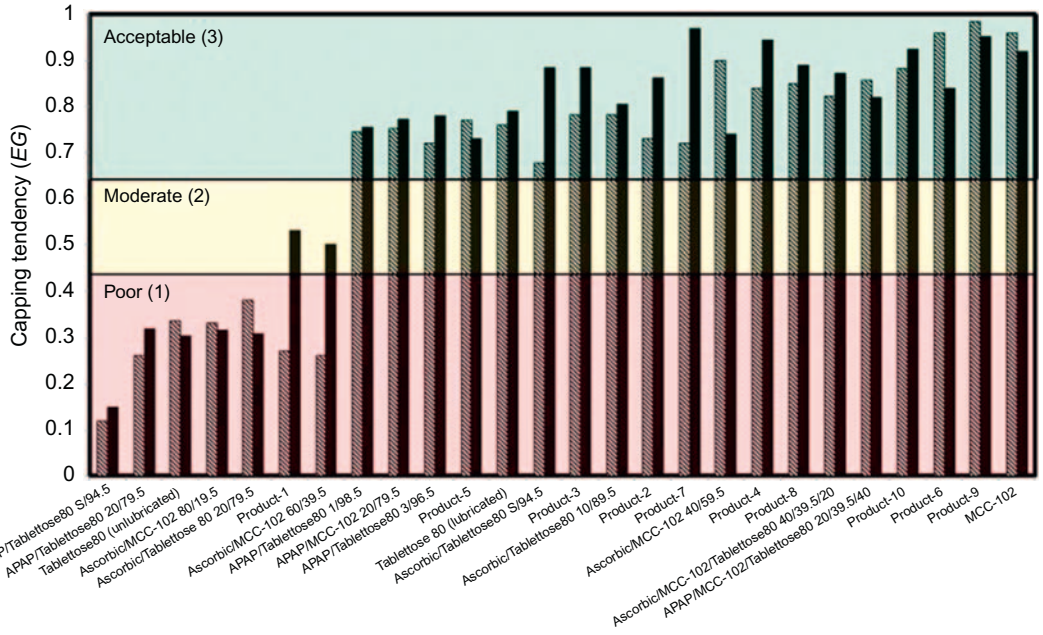


FIGURE 20.4 Nondestructive ultrasonic assessment of tablet capping tendency. Patterned and black colored bars are representing the $EG^{E} = E_{axial}/E_{radial}$ and $EG^{G} = G_{axial}/G_{radial}$, respectively.[71]

(MCC) can accommodate much higher drug load than brittle filler Tablettose 80. At 40% drug load, capping problems were absent with binary blends of ascorbic acid and MCC. At 60% drug load, a moderate capping issue was encountered. At 80% drug load, the risk for capping was significantly higher. These examples not only illustrate the need understand the impact of material properties on product manufacturability (ie, capping), but also the importance of selecting suitable excipients during formulation design to

accommodate undesirable material properties of a drug substance.

Besides material properties, processing variables can also significantly impact tablet capping tendency. In a subsequent study, Akseli et al. systematically investigated the effect of processing conditions on the same sample set. A DOEs (D-optimal surface response design) was used to study the effect of three factors on tablet capping tendency, that is, tooling shape (round, bevel edge, oval shape), compression force (5, 10, 15, 20, 25 KN), and compression speed (25, 40, 80 rpm). In this work, capping risk criteria were refined to indicate different regions of the *EG* ratio: poor/capping region: $EG < 0.4$, high-risk region: $0.4 < EG < 0.5$, low-risk region: $0.5 < EG < 0.6$, acceptable region: $EG > 0.6$. The impact of processing conditions can be illustrated using commercial products 1 and 2 in Figs. 20.5 and 20.6. As shown in the figures, Product 1 has moderate capping risk and Product 2 shows no capping tendency. The observations made during compression studies using a Pressuer in the lab are consistent with production experience of the commercial products. Because Product 1 has inherent tendency to cap, it is especially important to understand how the processing variables may affect risk of tablet capping. As shown in the Fig. 20.5, tooling shape has the most pronounced effect on tablet capping tendency with oval shaped tooling > bevel edged tooling > round tooling. Oval shaped tablets have moderate to high risk of capping under all conditions studied. The bevel edged tablets performed slightly better than the oval tablets in terms of capping risk. However, all tablets made using bevel edged tooling still show moderate to severe capping risk under all conditions. The round tablets offer the best opportunity for minimizing capping risk. Tablets made using round tooling have acceptable capping risk where compression force is very low. In contrast to Product 1, material properties of Product 2 are more desirable with acceptable capping risk. The risk for capping is low within a much larger processing space where the process parameters can be varied. Similar to Product 1, a pronounced impact of tablet tooling shape on capping tendency was observed with the same rank order: oval shape > bevel edge > round tooling. The safe process region is the widest for round tooling and the smallest for oval shape tooling. It should be noted that even though Product-2 has excellent material properties from a capping risk perspective, tablet capping may still occur under certain process conditions.

The general learnings from this case study indicate (1) both material properties and processing variables can have a significant impact on tablet quality and manufacturing. Therefore, it is essential to evaluate the potential influence of material properties, process variables, and their interplay on product quality and manufacturing during the design and development of product and process; (2) for formulations with material properties prone to have processing risks, options to minimize or eliminate the risk by varying process conditions are often limited. For formulations with improved material properties, process understanding is still crucial in developing a robust manufacturing process; (3) fast and material sparing predictive tools

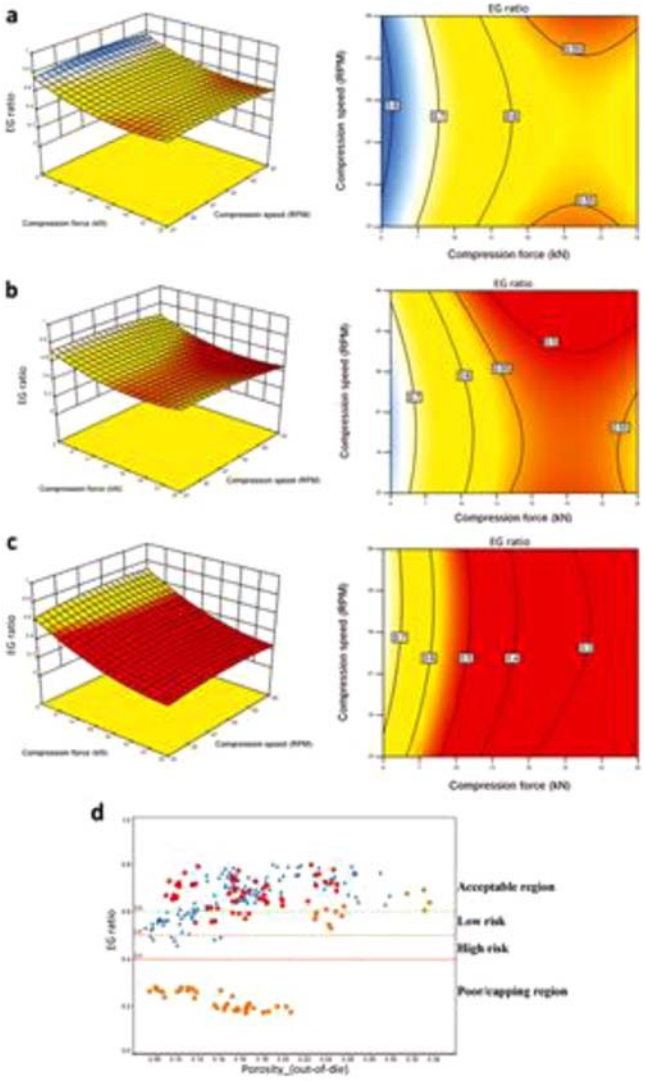


FIGURE 20.5 Three-dimensional surface plots and two-dimensional contours of *EG* ratio versus compression force versus compression speed for round (a), beveled edge (b), and oval shape tablets (c) manufactured using Product-1 blend. *EG* ratio as a function of tablet porosity for Product-1 (d) manufactured with round (circle), beveled edge (triangle), and oval shaped (square) toolings—poor/capping region: $EG < 0.4$, high-risk region: $0.4 < EG < 0.5$, low-risk region: $0.5 < EG < 0.6$, acceptable region: $EG > 0.6$.[72]

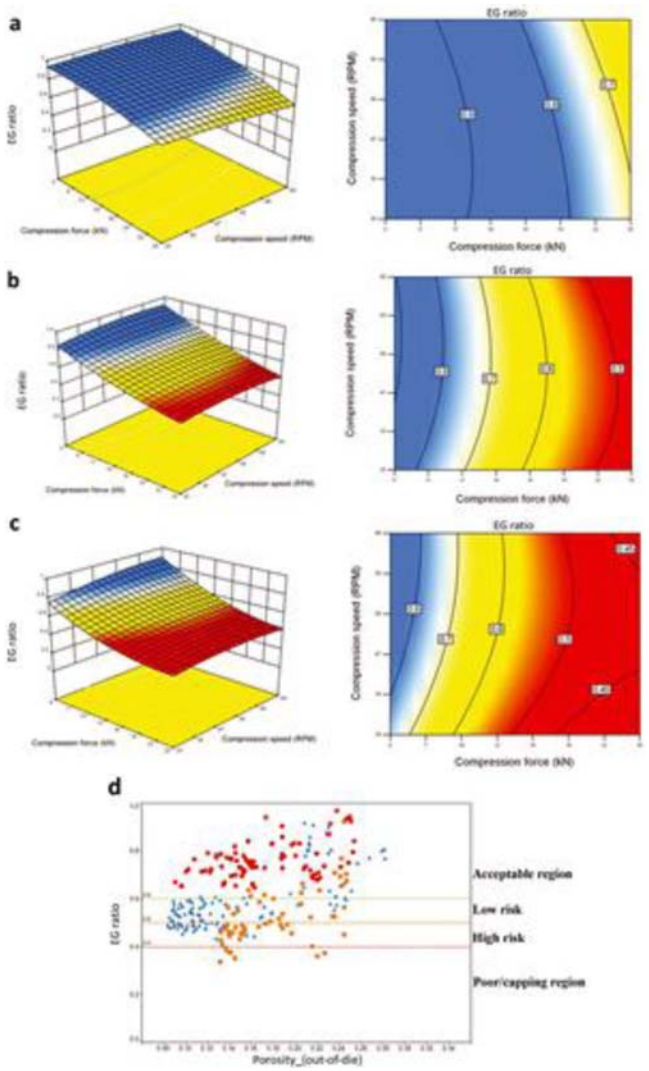


FIGURE 20.6 Three-dimensional surface plots and two-dimensional contours of *EG* ratio versus compression force versus compression speed for round (a), beveled edge (b), and oval shape tablets (c) manufactured using Product-1 blend. *EG* ratio as a function of tablet porosity for Product-2 (d) manufactured with round (circle), beveled edge (triangle), and oval shaped (square) tooling—poor/capping region: $EG < 0.4$, high-risk region: $0.4 < EG < 0.5$, low-risk region: $0.5 < EG < 0.6$, acceptable region: $EG > 0.6$.[72]

such as the ultrasonic capping test can be used to evaluate and anticipate risk early on, facilitate product and process development with enhanced understanding of CMAs and process parameters, and to develop a risk mitigation strategy.

20.4.2 Understanding formulation design of ER dosage forms of verapamil

Verapamil is an L-type calcium channel blocker used in the treatment of hypertension, angina pectoris, cardiac arrhythmia. Verapamil is a weak base with pKa of 8.8 and pH-dependent solubility (about 65 mg/mL at pH 1.2 and 4.5, about 13 mg/mL at pH 6.8). It is a BCS Class I compound with high permeability in upper and lower bowels.[74,75] The strengths of IR dosage forms range from 40 to 120 mg. The favorable dose, solubility, and permeability indicate that all three types of ER systems are suitable for this compound. This is reflected in the commercially available once-daily oral ER products which include a matrix (Isoptin SR and Calan SR tablets), reservoir (Verelan and Verelan PM capsules), and osmotic pump systems (Covera-HS tablets).

ER products of verapamil were first developed using hydrophilic matrix systems. Isoptin SR and Calan SR are matrix tablets based on sodium alginate, a pH-dependent, natural water-soluble linear unbranched polysaccharides consisting of different proportions of β-D-mannuronic acid (M) and α-L-guluronic acid (G) units.[76] At acidic pH below the pK_a of M (3.38) and G (3.65) monomers, sodium alginate is converted to insoluble and nonswellable alginic acid. In addition, solubility of the API increases with decreasing pH. Hence, drug release from the matrix tablets is dependent on pH as shown in Fig. 20.7a.[77] At pH 1.2, the in vitro release from matrix tablets is slow and mainly driven by diffusion because of high solubility of the API and the formation of the inert alginic acid. In addition, surface area available for drug release also changes over time due to tablet cracking and lamination at acidic pH.[78] At pH 4.5 and 6.8, sodium alginate hydrates and swells rapidly, forming a gel layer that controls the drug release rate. The data in Fig. 20.7a show that drug release at pH 4.5 was essentially governed by polymer erosion as indicated by the release kinetics in spite of high drug solubility. In addition, the drug release profile is essentially the same as that at pH 6.8 even though drug solubility is five fold lower at high pH. These observations can be attributed to the interaction between the protonated tertiary amine cation of verapamil and the carboxyl anion of sodium alginate, resulting in erosion-controlled drug release and a decrease in free drug concentration in the gel.

One of the shortcomings of both Isoptin SR and Calan SR tablets is that they need to be taken with food to maintain a narrow peak-to-trough ratio of the plasma profile. The significant food effect on oral absorption was believed to be, at least in part, associated with the effect of pH on drug release To overcome pH dependence of drug release, Howard and Timmins designed a new ER matrix system of verapamil by incorporating a pH-independent polymer in the alginate-containing formulation.[79] In an acidic environment (eg, stomach), the pH-independent polymer (eg, HPMC) hydrates to form a gel layer and serve to the

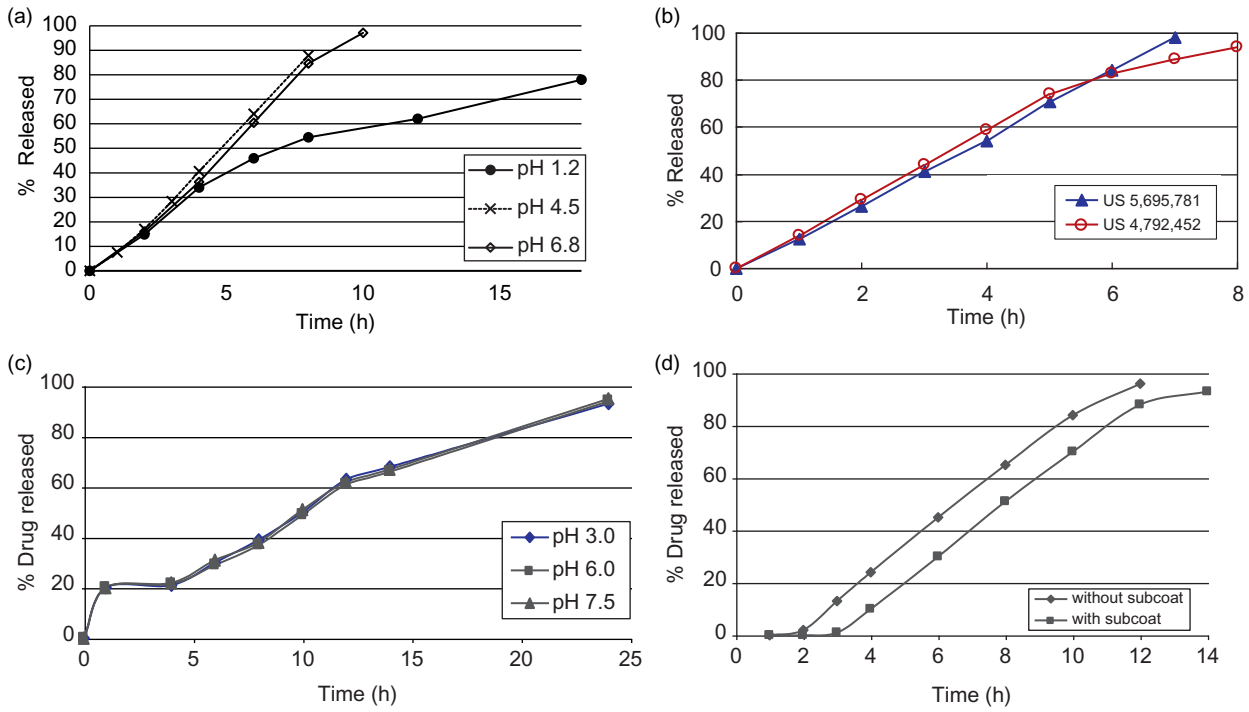


FIGURE 20.7 Drug release of verapamil from different ER systems: (a) pH-dependent matrix; (b) pH-independent matrices; (c) reservoir; and (d) osmotic pump.

control release rate of soluble API. As the pH increases with the passage of the tablets from the stomach to the intestinal tract, the alginate in the tablets starts to hydrate and form a gel layer that helps maintain lower pH in the matrix and contributes to modulation of the release rate. As a result, drug release from this system is independent of pH and follows zero-order kinetics for up to 80% of the dose as shown in Fig. 20.7b. To design a more robust pH-independent zero-order ER system, Zhang and Pinnamaraju incorporated an additional anionic polymer, a methacrylic acid copolymer, into the matrix system consisting of sodium alginate and a pH-independent polymer, HPMC.[80] Data in Fig. 20.7b showed that pH-independent zero-order release was obtained for up to nearly 100% of drug release. In both designs, the anionic methacrylate and/or alginate copolymers present in the HPMC-based matrix play a significant role in achieving pH-independent and zero-order release. The decreasing drug release rate over time from the matrix at higher pH due to decreasing solubility and surface area/diffusion path-length is compensated for by (1) maintaining lower gel layer pH and (2) interaction between the tertiary amine of the API and carboxyl group discussed previously.

ER capsule of verapamil hydrochloride, Verelan, is an ER reservoir system that contains a mixture of rapid and ER coated beads.[81] The formulation of the ER core consists of nonpareil seeds, API, a binder, talc, and a pH-modifier, malic acid with two pKas (3.4 and 5.1) to help achieve pH-independent release by providing low pH (high API solubility), hence more constant driving force for diffusion after being coated using the rate-controlling ethylcellulose film. The finished capsule dosage form contains 20% of the uncoated beads and 80% of the membrane-coated beads. Drug release profiles in Fig. 20.7c indicate that the drug release from this system is independent of pH. It should be noted that the original product was developed over two decades ago when the film coating polymers were applied in an organic solution. Significant advances have since been made in a polymer coating of solid dosage forms. Today, the majority of the membrane coating of a reservoir system is applied as aqueous dispersion that forms a film by a process of coalescence of submicron polymer particles. The film coating process is typically performed in a fluidized bed with Wurster insert. The properties of the resulting functional coating are influenced by coating formulations as well as processing variables. The details and in depth discussions of the coating process development are addressed in chapter "Development, Optimization, and Scale-Up of Process Parameters: Wurster Coating" of this book. To further improve the clinical outcome, the second generation reservoir ER system of verapamil, Verelan PM, was developed and approved in 1999. It was designed by applying chronotherapeutic principles to drug delivery,

or targeting drug release to the circadian rhythm of the disease state, to improve clinical responses.[82] Verelan PM is a bead-filled capsule that uses the CODAS (Chronotherapeutic Oral Drug Absorption System) ER technology designed for bedtime dosing by incorporating a 4- to 5-hour delay in drug delivery. The delay is controlled by the type, composition and level of release-modifying polymers applied to drug loaded beads. The drug release is essentially independent of pH, food intake, and GI motility.

Covera-HS tablet is also a chronotherapeutic ER dosage form of verapamil specifically designed to deliver the drug in higher concentrations during time of greatest need (eg, postawakening period) and in lesser concentrations at night when heart rate and blood pressure are lowest and cardiovascular events are least likely to occur. It is a two-chamber OROS Push-Pull osmotic system that consists of an active drug layer and an osmotic push layer. The device is designed to provide delayed onset of drug release by adding a coating layer between the tablet core and semipermeable outer membrane. Depending on the composition and level of the subcoat, the onset of drug release can be controlled. The ER drug delivery is independent of pH, GI motility, and food. This type of ER osmotic device is prepared with the following steps according to the issued US patents.[83–85]

(a) Composition and preparation of active drug layer:

Verapamil hydrochloride	600 g
Poly(ethylene oxide)	305 g
Sodium chloride	40 g
Polyvinylpyrrilidone	50 g
Magnesium stearate	5 g

All ingredients except magnesium stearate were granulated with anhydrous ethanol. The dried granules were lubricated with magnesium stearate. This procedure provides granules for the active drug layer.

(b) Composition and preparation of osmotic push layer:

Polyethylene oxide	735 g
Sodium chloride	200 g
Hydroxypropyl methylcellulose (Methocel E5)	50 g
Ferric oxide	10 g
Magnesium stearate	5 g

All ingredients except magnesium stearate were granulated with anhydrous ethanol. The dried granules were lubricated with magnesium stearate. This procedure provides granules for the osmotic push layer.

(c) Preparation of core tablet: A bilayer core tablet comprising an active layer and a push layer was compressed in a bilayer tablet press. A subcoat is applied on to the bilayer core tablets using enteric coating material.

It should be pointed out that manufacture of layered tablets often faces challenges as this type of tablets are prone to delaminate along the interface, causing layer separation during processing, scale-up or storage before coating, particularly when formulation and process are developed in isolation. Thus, to design a physically robust bilayer tablet with well-adhered layers, prospective assessment of delamination tendency of different layers of different compositions should be conducted through integrated evaluation of material properties, formulation, and process variables. Knowing the interaction, deformation and bonding properties between different materials during compaction is of vital importance to better understand the compaction behavior and possible failure mechanisms of bilayer tablets. For example, the deformation, relaxation, and bonding characteristics of a push layer primarily consisting of polyethylene oxide[86] should be evaluated if the active layer is predominantly deformed by brittle fracture. Significant mismatch of the physical and mechanical properties between the two layers often leads to difficulties in the development and manufacture of the layered tablets. Other parameters important to the development of layered tablets include level and extent of lubrication, granulation characteristics, tamping and compression forces, turret speed, and equipment.[87–89]

(d) Composition and preparation of semipermeable membrane:

Cellulose acetate	55%
Hydroxypropylcellulose	40%
Polyethylene glycol	5%

All ingredients in the formulation are dissolved in 80% acetone and 20% methanol. The bilayer tablets were coated in a pan coater. Two orifices were drilled on the side of the tablet containing the active drug. An optional color or clear coat may also be applied. It is noted that the most significant drawback of developing osmotic systems is the need for organic solvents in the manufacturing process. Others include tedious and complex processing operations that require extensive process studies and control.

Fig. 20.7d shows the in vitro drug release profiles of verapamil from the osmotic pump system. The dosage

form without and with an enteric subcoat exhibited a lag time of 1.5 and 3.0 hours in drug release, respectively, followed by zero-order release of verapamil.

20.4.3 Improving process robustness and capability through enhanced process understanding

FBG is a commonly employed unit operation in the pharmaceutical industry to improve flow properties and compressibility of fine powders by enlarging particle size.[90–93] The fluid-bed wet granulation process usually consists of three steps: dry mixing, spray agglomeration, and drying to the desired moisture level (as shown in Fig. 20.8).[90] Physical properties of granulations are dependent upon both formulation and operational variables that are associated with the process. For example, process humidity is known to be an important factor affecting granulation characteristics, thus, controlling the bed humidity is critical as it has a substantial impact on physical properties of the granules required for downstream processing including flow, packing efficiency, and compactibility.[90–93]

In order to improve process understanding and robustness of a legacy product, Lourenco et al. used in-line microwave resonance technology (MRT) to monitor granule moisture, temperature, and density of an industrial-scale FBG process with a batch size of 500 kg.[90] The commercial formulation used in the granulation study consisted of an API, a filler, and a binder which was applied as an aqueous solution. Through analysis of 192 batches using a combination of off-line particle size measurement and multivariate data analysis (MVDA), including multiway partial least squares (PLS), and principal component analysis (PCA), a significant seasonality effect on granule properties was observed. The multivariate batch control charts generated using the first principal component scores of the PLC model (Fig. 20.9a) show that process variability decreased over time during the mixing phase and eventually reached similar level for all production batches, indicating that the powder mixture reached homogeneity despite different initial state. However, for both the spraying and the drying phases (Fig. 20.9b and c), the batches produced from January to March (colder and drier months) are different from those made from April to July (warmer and wetter months). A much higher variability was noted at the end of spaying phase although the process parameters are controlled at the nominal conditions. This heterogeneity then led to the time profile shift in the drying phase, that is, longer drying time is generally required for the batches manufactured in warmer months when the humidity is higher. This seasonality effect was also confirmed by temperature and moisture profiles of the granules and the off-line particle size measurement (see Fig. 20.10). Compared to granules manufactured in the warmer months, the granules produced in the colder months had a lower moisture content, broader particle size distribution, and a higher portion of fines. Since all batches are manufactured with the same process parameters, observed high variability of this product was attributed to the influence of the inlet air relative humidity.

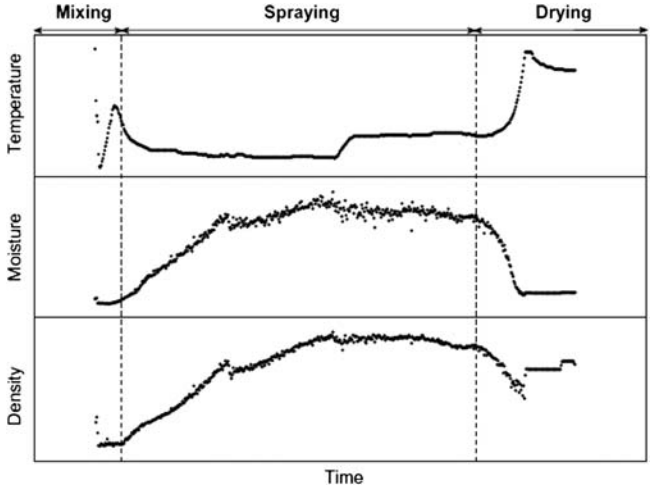


FIGURE 20.8 Typical time profiles of granule temperature, moisture, and density during fluid-bed granulation (FBG).[90]

Granules with highly variable property are undesirable because it can negatively impact the downstream process, resulting in poor process capability, manufacturing difficulty or even product failures. Low process capability due to high variability is not uncommon for many legacy products. This is because those legacy products are often manufactured under predetermined and fixed operating conditions defined without sufficient product and process understanding during development. Based on the understanding of the root cause for this product, one of the obvious approaches to reduce process variability would be to control inlet air humidity. However, it requires facility modification (such as an HAVC system) and equipment upgrades which are undesirable from a practical and economic standpoint. Thus, process optimization through enhanced process understanding was used as an alternative approach to decrease the intrinsic variability and improve process reproducibility. More specifically, the author conducted process re-development/optimization experiments by applying QbD approach.[91] All potential risks were identified and analyzed based on multidisciplinary risk assessment; three process parameters, including (1) inlet air temperature, (2) binder spray rate, and (2) air flowrate, were considered critical based on the relevant process

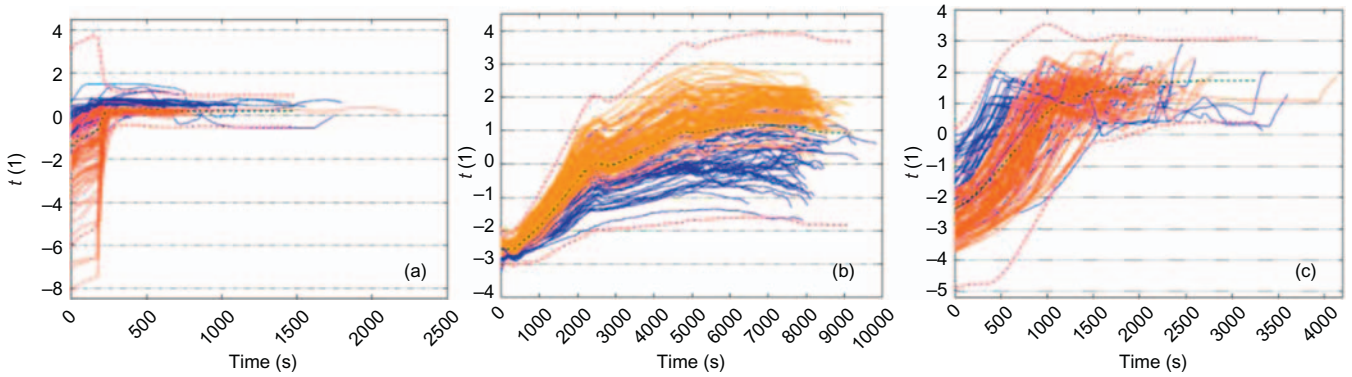


FIGURE 20.9 Multivariate batch control charts for (a) mixing phase, (b) spraying phase, and (c) drying phase. (Batches are colored according to production month (January–March: black, April–July: grey).)[90]

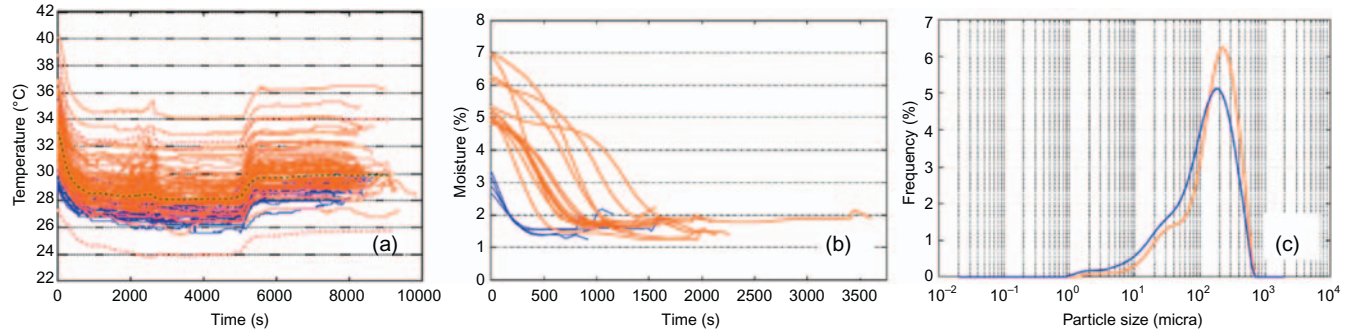


FIGURE 20.10 (a) Granule temperature profile during spraying phase, (b) granule moisture profiles during drying phase, and (c) average granule particle size distribution. (Batches are colored according to production month (January–March: black, April–July: grey).)[90]

knowledge (empirical and theoretical). Two sets of DOEs, a screening DOE and an optimization DOE, were executed on the pilot scale with a batch size of approximate 180 kg. In general, an optimization DOE should be performed within the same space as the screening DOE. However, results from the two DOEs were combined and analyzed simultaneously using PCA in this study because the binder spray rate was intentionally altered outside the range explored in the screening DOE (Fig. 20.11). The effect of CPPs on CQAs of the granules (ie, particle size distribution (PSD), bulk and tapped densities, flowability, and angle of repose) was investigated. The results demonstrated that the granule quality is strongly correlated with the moisture profile during the spraying phase. The granules manufactured under wetter conditions have improved flowability, narrower PSD, and less agglomerates (ie, smaller impact by drying phase conditions).

Taking into consideration the multivariate interactions, a design space was determined using a mathematical model between the CPPs and CQAs and shown in Fig. 20.12. In Fig. 20.12a, the desired dimensional space corresponds to the granules flowing through an orifice with values lower than 5.5 s/100 g as marked by a filled circle. The design space of the granule moisture time profile can be computed using the model as shown in Fig. 20.12b. Outside the design space, there is a risk of either a fluidization issue/bed collapse (above the upper limit) or decreased granule quality (below the lower limit). The corresponding design space for the inlet air temperature and binder spray rate in the spraying phase are shown in Fig. 20.12c, which excludes the lower right corner where granules flow through an orifice for which the value is equal or higher than 5.5. Density and flow properties of the pilot-scale granules manufactured within the design space showed lower variability compared to the historical commercial-scale granules (Fig. 20.13). A narrower PSD with a decrease in fine particles was also observed (Fig. 20.14). These granules were successfully compressed into tablets that meet all quality requirements.

To apply the study findings to the commercial-scale FBG process, acceptable limits of critical parameters were calculated using the scale-up approach of matching (1) solid density in FBG, (2) air flow Reynolds number, and (3) a process time constant related to spraying flux. A comparison between the newly proposed process ranges and the existing filed limits and

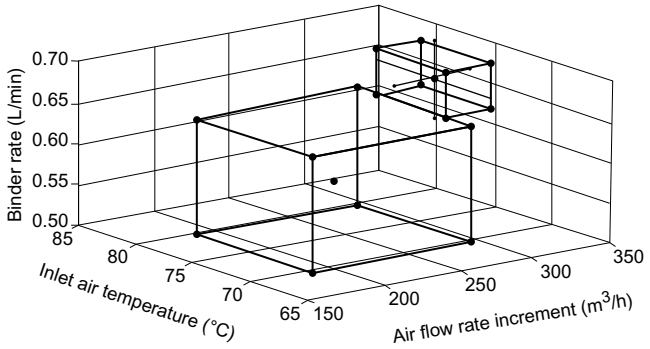


FIGURE 20.11 Pilot-scale design of experiments (Screening DOE: squares, optimization DOE: circles and stars).[91]

operating ranges is provided in Fig. 20.15. The original process consists of five spraying subphases (ie, the five white bars for air flowrate). Since the air flow rate was proven to be a noncritical process parameter for granule CQAs based on pilot-scale DOEs, a simpler process consisting of only one set of flow rate conditions is proposed.

Based on enhanced process understanding, granules quality variability due to seasonality effect can be decreased using an in-line moisture analyzer in combination with the mathematical model that correlates CPPs with moisture content and other quality attributes of the granules as control strategies. This approach offers process flexibility while accommodating the inlet air humidity variability. However, if such a strategy could not be implemented, the seasonality effect can be reduced or minimized by applying the

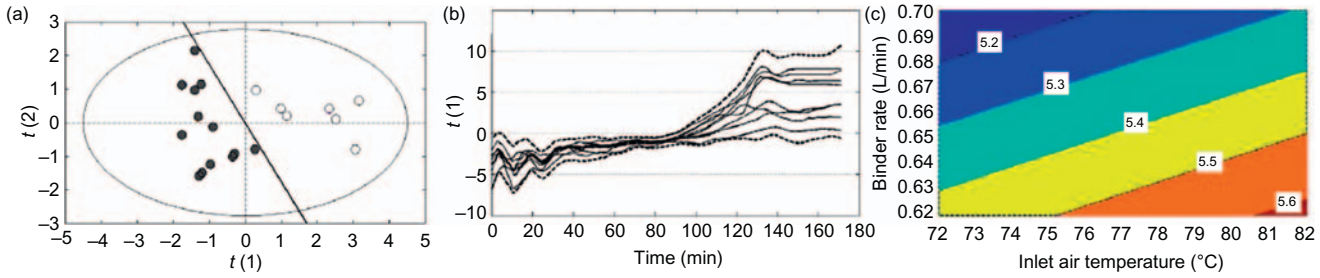


FIGURE 20.12 (a) Design space for granule attributes in the low dimensional space ($t1$ vs $t2$) of their mechanical properties, (b) design space for the granules moisture profile defined in the original variable, (c) design space for the critical process parameters (at the pilot scale).[91]

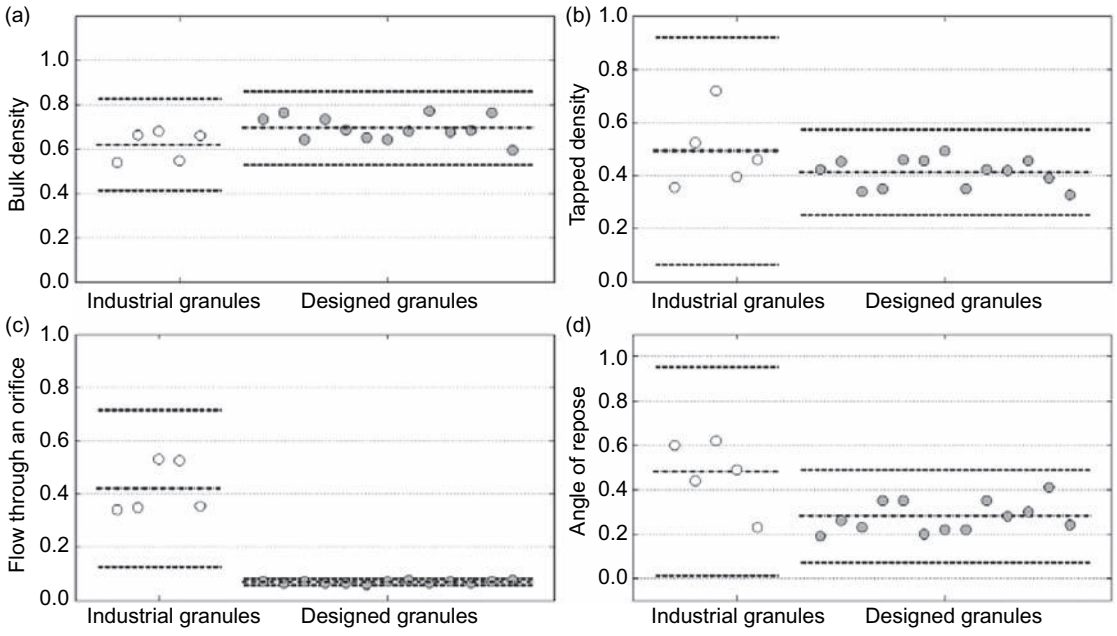


FIGURE 20.13 Physical properties comparison of pilot-scale granules generated within design space (filled circle) versus commercial-scale granules (empty circle).[91]

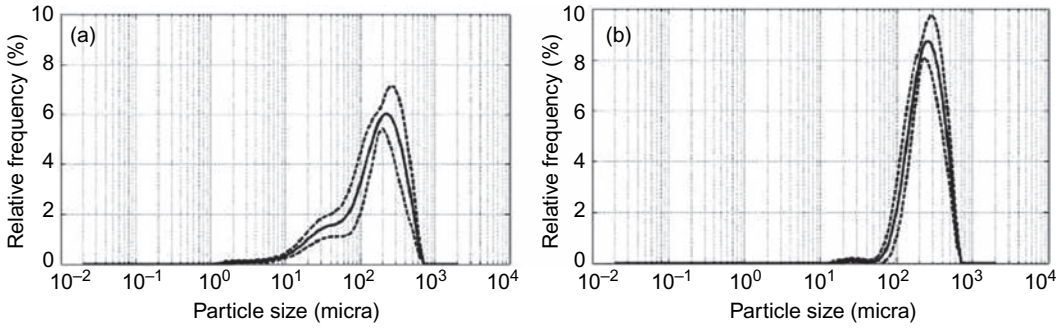


FIGURE 20.14 PSD comparison of pilot-scale granules generated within design space (filled circle) versus commercial-scale granules (empty circle).[91]

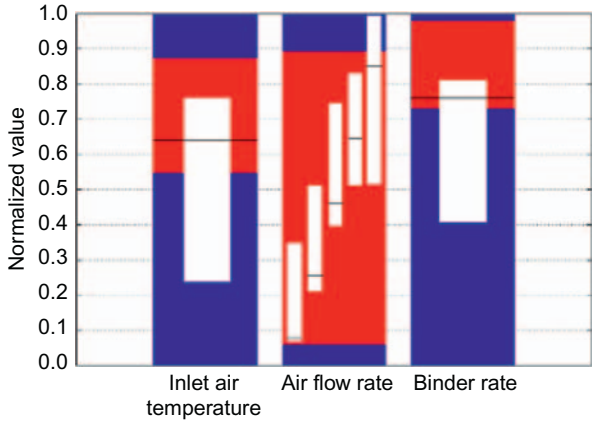


FIGURE 20.15 Range and set points for inlet air temperature, airflow rate, and binder spray rate (submitted limits: black, nominal operational range: white, proposed design space: grey).[91]

newly defined design space. Since the drying time profile shift shown in Fig. 20.9c is due to the seasonality effect in the spraying phase, the focus needs to be placed on consistently achieving sufficient granule moisture content during the spraying/granulation phase. Based on the mass balance, the final granule moisture is determined by the water introduced into the system and the moisture removed by airflow. The total water is introduced into the system via two routes: (1) binder solution which is normally well controlled and (2) moisture present in the inlet air which varies with season and is identified as the primary cause for the observed variable granule quality. Decreasing the portion of inlet air moisture presented in total water will translate to lower impact from the seasonality effect. Compared with the granule moisture profiles of the existing commercial-scale process, significant higher granule moisture with much tighter control range is proposed for the improved process (as shown in Fig. 20.16) and the higher granule moisture can be achieved by adjusting the controlled process parameters include increasing binder spray rate, decreasing air flowrate to lower the drying capacity, or the combination of both. This study shows that the seasonality effect on granule quality due to uncontrolled moisture brought into the system from the inlet air can be reduced to improve process capability through enhanced product and process understanding.

20.5 INTELLECTUAL PROPERTY CONSIDERATIONS

Intellectual property (IP) rights are central to the pharmaceutical industry because of the high R&D costs, low success rate, and lengthy product testing required before a drug product can enter the market. In general, IP estate of pharmaceuticals not only encompasses patents that protect innovations, but also regulatory or legislative exclusivity, trade secrets that are difficult to reverse-engineer. Among various types of patents, the most important are the composition-of-matter (ie, chemical entity) patents. However, the creation of an IP portfolio to cover a range of drug delivery characteristics and related dosage forms has also become a key strategic element in extending product life cycles. Life cycle management can often play a significant role by offering differentiated products with clinical benefits when the new product development, medical need, and market requirements are aligned. Over the past few decades, many companies have reaped great benefits from this type of IP generation. The patentable subject matter (or statutory subject matter) in dosage form development includes modifications of structure (eg, new solid phases, salts), formulation (eg, composition, technology, particle size or surface area of the API), method of treatment (eg, route of delivery, dosing regimen), process, methods of improving bioavailability or overcoming stability problems, newly discovered clinical benefits (eg, decreased adverse events, new indications), or performance (drug release profiles, PK or PD outcomes) resulting from MR delivery independent of delivery technologies applied.[94] One of the classic examples

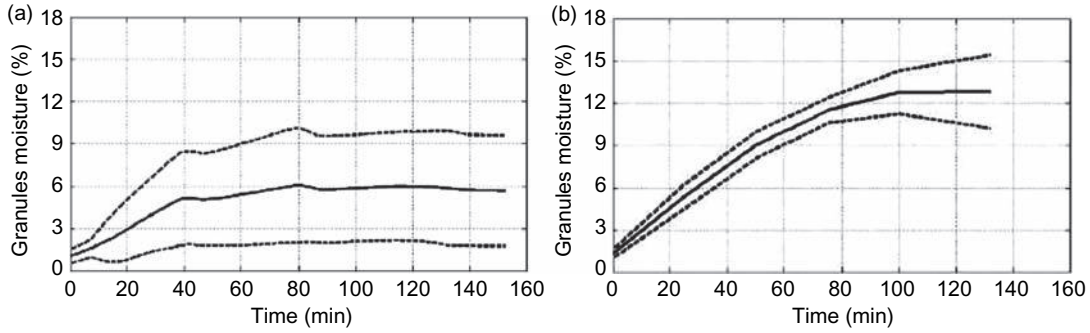


FIGURE 20.16 Granule moisture profiles during the spraying phase of (a) commercial-scale granules, (b) proposed granules.[91]

involves development of patents protecting niche products of nifedipine (Procardia XL and Adalat CC). Because the 3-year regulatory exclusivity expired in 1992, the ER products had to rely on the exclusivity provided by patents that claim benefits of ER delivery and controlled particle size for enhancing bioavailability. Among four patents listed in the "Orange Book,"[95] the most critical is US patent 5,264,446 that claims, "A solid pharmaceutical composition comprising as the active ingredient an effective amount of nifedipine crystals with a specific surface area of 1.0−4.0 m^2/g, in admixture with a solid diluent, to result in a sustained release of nifedipine." As a result, the innovator product retained exclusivity for an additional 9 years until a competitor utilized finer particle size and amorphous drug, as disclosed in the US patent 5,455,046. Examples of more recent successful drug delivery/product patents for marketed products include: Nexium (esomeprazole magnesium), Allegra (fexofenadine), Tricor (fenofibrate), Concerta (methylphenidate), and Ambien CR (zolpidem tartrate).

It should be noted that most of the above mentioned products were developed at a time when there were fewer companies with novel drug delivery development expertise and experience. Because of the increased scientific know-how and technical maturity, as well as more crowded patent landscape in drug delivery, formulation and drug delivery technology-based patents have become more difficult to obtain or have lower hurdles to get around after granting. In protecting innovation in the manufacturing process or method, trade-secrets or unpatented proprietary know-how may also be considered because process patents can sometimes be difficult to enforce. One of the classical examples is the successful protection of commercial manufacturing know-how of stable paclitaxel protein-bound nanoparticles that consist of 90% albumin and 10% API. Therefore, an effective and robust IP strategy for a new drug product should be built upon continuous development and review of both in-house and in-licensed opportunities throughout product life cycle. More specifics and examples can be found in Chapters 39 and 40 in the first edition of this book.

20.6 SUMMARY

Development of a new product consists of a sequence of stages, beginning with an initial product concept or idea followed by design, testing, development, technology transfer, and commercialization. To effectively and efficiently turn an idea on paper into a high-quality solid dosage form that is demonstrable, producible and manufacturable, it is imperative to integrate formulation and process design and utilize a systematic approach throughout the development process. In most cases, quality and manufacturing challenges or problems encountered during pharmaceutical development or commercial production can be attributed to a design deficiency or flaws due to inadequate product and process understanding. Efforts spent in addressing or "patching" these problems at later stages of a product lifecycle not only are costly and time-consuming but are also often unable to resolve the underlying cause. This chapter emphasizes the importance of applying multidisciplinary knowledge for the fundamental understanding of drug substance, excipients, delivery technology that matches drug properties, processing technology, and their interplay to enable rational design and development of solid dosage forms. A high level of product and process understanding during the design and development phase is the key to ensure that (1) a product can meet its predefined quality attributes and (2) the manufacturing process is capable of consistently producing drug products of the same quality while accommodating the expected variability of raw materials, operating conditions, process equipment, environmental conditions, and human factors.

References

1. FDA Unveils New Initiative To Enhance Pharmaceutical Good Manufacturing Practices. *U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research*. <http://www.fda.gov/bbs/topics/NEWS/2002/NEW00829.html>; 2002 [21.08.2002].
2. Yu L. Pharmaceutical quality by design: product and process development, understanding and control. *Pharm Res* 2008;**25**(4):781–91.
3. Guidance for Industry, Q8 Pharmaceutical Development. U.S. Food and Drug Administration. May 2006.
4. Office of Generic Drugs White Paper: Question-Based Review (QbR) for Generic Drugs: An Enhanced Pharmaceutical Quality Assessment System. U.S. Food and Drug Administration. 2007.
5. Haleblian J, McCrone W. Pharmaceutical applications of polymorphism. *J Pharm Sci* 1969;**58**(8):911–29.
6. Aguiar AJ, Krc J, Kinkel AW, Samyn JC. Effect of polymorphism on the absorption of chloramphenicol from chloramphenicol palmitate. *J Pharm Sci* 1967;**56**(7):847–53.
7. Zhang GZ, Law D, Schmitt EA, Qiu Y. Phase transformation considerations during process development and manufacture of solid oral dosage forms. *Adv Drug Del Rev* 2004;**56**(3):371–90.
8. Zhou D, Qiu Y. Understanding drug properties in formulation and process design of solid oral products. *J Validation Tech* 2010;**16**(4):74–84.
9. US Pharmacopeia USP38/NF33, 2015. US Pharmacopeial Convetion, Rockville, MD, general chapter <905>, pp. 2778–2784.
10. Yalkowsky SH, Bolton S. Particle size and content uniformity. *Pharm Res* 1990;**7**(9):962–6.
11. Fan A, Parlerla S, Carlson G, Ladipo D, Dukich J, Capella R. Effect of particle size distribution and flow property of powder blend on tablet weight variation. *Am Pharm Rev* 2005;**8**:73–8.
12. Alderborn G, Nyström C. Studies on direct compression of tablets. IV. The effect of particle size on the mechanical strength of tablets. *Acta pharmaceutica Suecica* 1982;**19**(5):381–90.
13. Byrn SR, Pfeiffer RR, Stowell JG. *Solid state chemistry of drugs*. 2nd ed. West Lafayette, Indiana: SSCI; 1999.
14. Ruotsalainen M, Heinämäki J, Taipale K, Yliruusi J. Influence of the aqueous film coating process on the properties and stability of tablets containing a moisture-labile drug. *Pharm Dev Tech* 2003;**8**(4):443–51.
15. Nokhodchi A. An overview of the effect of moisture on compaction and compression. *Pharm Tech* 2005;**29**(1):46–66.
16. Amidon GL, Lennernäs H, Shah VP, Crison JR. A theoretical basis for a biopharmaceutical drug classification: the correlation of in vivo drug product dissolution and in vivo bioavailablity. *Pharm Res* 1995;**12**(3):413–20.
17. Benet L. The role of BCS (biopharmaceutics classification system) and BDDCS (biopharmaceutics drug disposition classification system) in drug development. *J Pharm Sci* 2013;**102**(1):34–42.
18. Lieberman HA, Lachman L, Schwartz JB. *Pharmaceutical dosage form: tablets*. 2nd ed. New York: Marcel Dekker; 1989.
19. Chen ML, Straughn AB, Sadrieh N, Meyer M, Faustino PJ, Ciavarella AB, et al. A modern view of excipient effects on bioequivalence: case study of sorbitol. *Pharm Res* 2007;**24**:73–80.
20. Chen ML, Sadrieh N, Yu L. Impact of osmotically active excipients on bioavailability and bioequivalence of BCS Class III drugs. *AAPS J* 2013;**15**(4):1043–50.
21. Parrot EL. Milling. In: Lachman L, Lieberman HA, Kanig JL, editors. *The theory and practice of industrial pharmacy*. 3rd ed. Philadelphia: Lea & Febiger; 1986. p. 21–46.
22. Buckton G. *Interfacial phenomena in drug delivery and targeting*. London: Harwood Academic; 1995. p. 29.
23. Hersey A, Robinson BH, Kelly HC. Mechanisms of inclusion-compound formation for binding of organic dyes, ions and surfactants to α-cyclodextrin studied by kinetic methods based on competition experiments. *J Chem Soc, Faraday Trans 1* 1986;**82**(5):1271–87.
24. Rajewski RA, Stella VJ. Pharmaceutical applications of cyclodextrins. 2. In vivo drug delivery. *J Pharm Sci* 1996;**85**(11):1142–69.
25. Thompson DO. Cyclodextrins – enabling excipients: their present and future use in pharmaceuticals. *Crit Rev Ther Drug Carrier Syst* 1997;**14**(1):1–104.
26. Uekama K, Hirayama F, Irie T. Applications of cyclodextrins in pharmaceutical preparations. *Drug Targeting Delivery* 1994;**3**:411–56.
27. Gao P, Morozowich W. Development of supersaturatable self-emulsifying drug delivery system formulations for improving the oral absorption of poorly soluble drugs. *Expert Opin Drug Delivery* 2006;**3**(1):97–110.
28. Gao P, Rush BD, Pfund WP, Huang T, Bauer JM, Morozowich W, et al. Development of a supersaturable SEDDS (S-SEDDS) formulation of paclitaxel with improved oral bioavailability. *J Pharm Sci* 2003;**92**(12):2386–98.
29. Simonelli AP, Mehta SC, Higuchi WI. Inhibition of sulfathiazole crystal growth by polyvinylpyrrolidone. *J Pharm Sci* 1970;**59**(5):633–8.
30. Sekikawa H, Fujiwara J, Naganuma T, Nakano M, Arita T. Dissolution behaviors and gastrointestinal absorption of phenytoin in phenytoin-polyvinylpyrrolidone coprecipitate. *Chem Pharm Bull* 1978;**26**(10):3033–9.
31. Megrab NA, Williams AC, Barry BW. Oestradiol permeation through human skin and silastic membrane: effects of propylene glycol and supersaturation. *J Control Release* 1995;**36**(3):277–94.
32. Raghavan SL, Trividic A, Davis AF, Hadgraft J. Crystallization of hydrocortisone acetate: influence of polymers. *Int J Pharm* 2001;**212**(2):13–221.
33. Raghavan SL, Kiepfer B, Davis AF, Kazarian SG, Hadgraft J. Membrane transport of hydrocortisone acetate from supersaturated solutions; the role of polymers. *Int J Pharm* 2001;**221**(1-2):105.
34. Suzuki H, Sunada H. Comparison of nicotiamide, ethylurea and polyeythylene glycol as carriers for nifedipine solid dispersion systems. *Chem Pharm Bull* 1997;**45**(10):1688–93.
35. Kohri N, Yamayoshi Y, Xin H, Iseki K, Sato N, Todo S, et al. Improving the oral bioavailability of albendazole in rabbits by the solid dispersion technique. *J Pharm Pharmacol* 1999;**51**(2):159–64.
36. Raghavan SL, Trividic A, Davis AF, Hadgraft J. Effects of cellulose polymers on supersaturation and in vitro membrane transport of hydrocortisone acetate. *Int J Pharm* 2000;**193**(2):231–7.
37. Iervolino M, Raghavan SL, Hadgraft J. Membrane penetration enhancement of ibuprofen using supersaturation. *Int J Pharm* 2000;**198**(2):229–38.
38. Gao P, Guyton ME, Huang T, Bauer JM, Stefanski KJ, Lu Q. Enhanced oral bioavailability of a poorly water soluble drug PNU-91325 by supersaturatable formulations. *Drug Dev Ind Pharm* 2004;**30**(2):221–9.
39. Suzuki H, Sunada H. Influence of water-soluble polymers on the dissolution of nifedipine solid dispersion with combined carriers. *Chem Pharm Bull* 1998;**46**(3):482–7.
40. Suzuki H, Miyamoto N, Masada T, Hayakawa E, Ito K. Solid dispersions of benidipine hydrochloride. II. Investigation of the interactions among drug, polymer and solvent in preparations. *Chem Pharm Bull* 1996;**44**(2):372–7.
41. Matsumoto T, Zografi G. Physical properties of solid molecular dispersions of indomethacin with poly(vinylpyrrolidone) and

poly(vinylpyrrolidone-co-vinyl-acetate) in relation to indomethacin crystallization. *Pharm Res* 1999;**16**(11):1722–8.
42. Pharmaceutical Benchmarking Forum, KMR Group data, 2012.
43. Wasylaschuk WR, Harmon PA, Wagner G, Harman AB, Templeton AC, Xu H, et al. Evaluation of hydroperoxides in common pharmaceutical excipients. *J Pharm Sci* 2007;**96**(1):106–16.
44. M.D. Ruff, et al. Pharmaceutical composition containing bupropion hydrochloride and a stabilizer. US Patent 5,358,970.
45. A. Maitra, et al. Pharmaceutical composition containing bupropion hydrochloride and an inorganic acid stabilizer. US Patent 5,968,553.
46. W. Oberegger, et al. Modified release formulations of a bupropion salt, US Patent Application #20070027213.
47. U.S. Government Accountability Office. Drug shortages: public health threat continues despite efforts to help ensure product availability. 2014. Retrieved February 12, 2014, from: <http://www.gao.gov/assets/670/660785.pdf>.
48. 3rd ed. Augsburger LL, Hoag SW, editors. *Pharmaceutical dosage forms—tablets*, vol. 1-3. Boca Raton, FL: CRC Press; 2008.
49. Wardrop J, Law D, Qiu Y, Engh K, Faitsch L, Ling C. Influence of solid phase and formulation processing on stability of ABT-232 tablet formulations. *J Pharm Sci* 2006;**95**(11):2380–92.
50. Tantry JS, Tank J, Suryanarayanan R. Processing-induced phase transitions of theophylline—implications on the dissolution of theophylline tablets. *J Pharm Sci* 2007;**96**(5):1434–44.
51. Antony J. *Design of experiments for engineers and scientists*. Burlington, MA: Elsevier Inc; 2003.
52. Hwang R, Kowalski DL. Design of experiments for formulation development. *Pharm Tech* 2005;**12**.
53. Hwang R, Peck GR, Besserman DM, Friedrich CE, Gemoules MK. Tablet relaxation and physicomechanical stability of lactose, microcrystalline cellulose,and dibasic calcium phosphate. *Pharm Tech* 2001;**11**:54–80.
54. Qiu Y, Garren J, Samara E, Abraham C, Cheskin HS, Engh KR. Once-a-day controlled-release dosage form of Divalproex Sodium II: development of a predictive in vitro release method. *J Pharm Sci* 2003;**92**(11):2317–25.
55. Nelloreb RV, Rekhia GS, Hussainc AS, Tillmand LG, Augsburgera LL. Development of metoprolol tartrate extended-release matrix tablet formulations for regulatory policy consideration. *J Control Release* 1998;**50**(1-3):247–56.
56. Draft Consensus Guideline: Pharmaceutical Development Annex To Q8. In: *International conference on harmonization of technical requirements for registration of pharmaceuticals for human use*; November, 2007.
57. Levin M, editor. *Pharmaceutical process scale-up*. 2nd ed. New York: Marcel Dekker, Inc; 2005.
58. Nash RA, Wachter AH. *Pharmaceutical process validation*. 3rd ed. New York: Marcel Dekker, Inc; 2005.
59. PQRI White Paper: Process Robustness, 2005.
60. Long CP, McQuaid J. Strategic approaches to process optimization and scale-up. *Pharm Technol* 2010;**34**(9).
61. Zlokarnik M. *Dimensional analysis and scale-up in chemical engineering*. Berlin-New York: Springer-Verlag; 1991.
62. Levin M. How to scale up scientifically. *Pharm Technol* 2005;**s4**.
63. Horsthuis GJB, et al. Studies on upscaling parameters of the Gral high shear granulation process. *Int J Pharm* 1993;**92**(143).
64. Pandey P, Badawy SF. A quality by design approach to scale-up of high-shear wet granulation process. *Drug Dev Ind Pharm* 2016;**42**(2):1–15.
65. U.S. Food and Drug Administration. Guidance for industry: PAT—a framework for innovative pharmaceutical development, manufacturing, and quality assurance, 2004.
66. U.S. FDA Guidance for Industry on Process Validation: general principles and practices, 2011.
67. European Medicines Agency. Guideline on process validation for finished products—information and data to be provided in regulatory submissions, 2014.
68. Khlebnikova E. Statistical tools for process qualification. *J Validation Technol* 2012;**18**(2):93–9.
69. ISPE Good Practice Guide: technology transfer (2nd ed.), 2014.
70. WHO Technical Report Series, No. 961. Annex 7 WHO guidelines on transfer of technology in pharmaceutical manufacturing, 2011.
71. Akseli I, Ladyzhynsky N, Katz J, He X. Development of predictive tools to assess capping tendency of tablet formulations. *Powder Technol* 2013;**236**:139–48.
72. Akseli I, Stecuła A, He X, Ladyzhynsky N. Quantitative correlation of the effect of process conditions on the capping tendencies of tablet formulations. *J Pharm Sci* 2014;**103**(6):1652–63.
73. Garekani HA, Ford JL, Rubinstein MH, Rajabi-Siahboomi AR. Effect of compression Force, Compression Speed, and Particle Size on the Compression Properties of Paracetamol. *Drug Dev Ind Pharm* 2001;**27**(9):935–42.
74. Vogelpoel H, Welink J, Amidon GL, Junginger HE, Midha KK, Möller H, et al. Biowaiver monographs for immediate release solid oral dosage forms based on biopharmaceutics classification system (BCS) literature data: verapamil hydrochloride, propranolol hydrochloride, and atenolol. *J Pharm Sci* 2004;**93**:1945–56.
75. Guidance for Industry: waiver of in vivo bioavailability and bioequivalence studies for immediate-release solid oral dosage forms based on a biopharmaceutics classification system. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research. 2000.
76. Chan LW, Ching AL, Liew CV, Heng PWS. Mechanistic study on hydration and drug release behavior of sodium alginate compacts. *Drug Dev Ind Pharm* 2007;**33**(6):667–76.
77. Lin Z, Zhou D, Hoag S, Qiu Y. Influence of drug properties and formulation on in vitro drug release and biowaiver regulation of oral extended release dosage forms. *AAPS J* 2016;**18**(2):333–45.
78. Mandal S, Basu S, Sa B. Sustained release of a water-soluble drug from alginate matrix tablets prepared by wet granulation method. *AAPS PharmSciTech* 2009;**10**(4):1348–56.
79. Howard JR Timmins P. Controlled release formulation, US Patent 4,792,452; 1988.
80. Zhang G Pinnamaraju P. Sustained release formulation containing three different types of polymers, US Patent 5,695,781; 1997.
81. Panoz DE Geoghegan EJ. Controlled absorption pharmaceutical composition, US Patent 4,863,742; 1989.
82. Sica DA, White W. Chronotherapeutics and its role in the treatment of hypertension and cardiovascular disease. *J Clin Hypertens* 2000;**2**(4):279–86.
83. Jao F, Wong PS, Huynh HT, McChesney K, Wat PK. Verapamil therapy, US Patent 5,160,744; 1992.
84. Jao F, Wong PS, Huynh HT, McChesney K, Wat PK. Verapamil therapy, US Patent 5,190,765; 1993.
85. Jao F, Wong PS, Huynh HT, McChesney K, Wat PK. Verapamil therapy, US Patent 5,252,338; 1993.
86. Yang L, Venkatesh G, Fassihi R. Characterization of compressibility and compactibility of poly(ethylene oxide) polymers for modified release application by compaction simulator. *J Pharm Sci* 1996;**85**(10):1085–90.
87. Podczecka F, Drakeb KR, Newtonb JM, Haririanc I. The strength of bilayered tablets. *Eur J Pharm Sci* 2006;**29**(5):361–6.
88. Yang L, Venkatesh G, Fassihi R. Compaction simulator study of a novel triple-layer tablet matrix for industrial tableting. *Inter J Pharm* 1997;**152**(1):45–52.

89. Li SP, Karth MG, Feld KM, Di Paolo LC, Pendharkar CM, Williams RO. Evaluation of bilayer tablet machines: a case study. *Drug Dev Ind Pharm* 1995;**21**(5):571–90.
90. Louenco V, Herdling T, Reich G, Menezes J, Lochmann D. Combining microwave resonance technology on multivariate data analysis as a novel pat tool to improve process understanding in fluid bed granulation. *Eur J Pharm Biopharm* 2011;**78**:513–21.
91. Louenco V, Lochmann D, Reich G, Menezes J, Herdling T, Schewitz J. A quality by design study applied to an industrial pharmaceutical fluid bed granulation. *Eur J Pharm Biopharm* 2012;**81**:438–47.
92. Chauhury A, Niziolek A, Ramachandran R. Multi-dimensional mechanistic modeling of fluid bed granulation processes: an integrated approach. *Adv Powder Technol* 2013;**24**:113–31.
93. Faure A, York P, Rowe R. Process control and scale-up of pharmaceutical wet granulation processes: a review. *Eur J Pharm Biopharm* 2001;**52**:269–77.
94. Qiu Y. Intellectual property considerations in pharmaceutical development: basics and examples. *CPA/AAPS short course: modern theory and practice of pharmaceutical sciences in oral dosage form R&D*. Shanghai, June 6–10, 2005.
95. Approved Drug Products with Therapeutic Equivalence Evaluations. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research, Office of Pharmaceutical Science, Office of Generic Drugs; <http://www.accessdata.fda.gov/scripts/cder/ob/>.

This page intentionally left blank

CHAPTER

21

Analytical Development and Validation for Solid Oral Dosage Forms

X.(F.) Fang[1], *G. Carr*[2] *and R.C. Freeze*[3]

[1]Novast Laboratories Ltd, Nantong, China [2]Patheon Inc., Mississauga, ON, Canada
[3]Regulatory Affairs Abbott Laboratories, Abbott Park, IL, United States

To ensure the quality of active pharmaceutical ingredients (APIs) and finished drug products, analytical methods must be developed and validated for the intended use. Method development and validation for submission for marketing application and for life cycle management of products is clearly outlined in guidance documents published by the agencies such as the International Conference on Harmonization (ICH)[1] and Food and Drug Administration (FDA).[2] However, the requirements of method development and validation during the earlier phases of product development vary, and this is a worthy topic for discussion. In addition to the many guidance documents that discuss method validation, there are several books available that provide a detailed interpretation of the existing regulations and discuss the appropriate techniques for method development and validation.[3,4] This chapter outlines a strategic approach to the development and validation of analytical procedures for solid oral dosage forms from early development, registration of marketing authorization, to postapproval of market products. With this information, as well as by giving details of case studies, we intend to aid readers in understanding the required aspects of method validation and method transfers.

21.1 ANALYTICAL METHOD DEVELOPMENT AND VALIDATION STRATEGY

A strategic approach to the development and validation of analytical methods is critical for the efficient operation of a product development group. This strategic approach balances the amount of validation required to ensure that a method is appropriately validated for the phases of product development with the need to be cost effective and meet tight project timelines.

Existing regulations allow a great deal of flexibility in terms of the amount of data that needs to be submitted during the course of drug development. For clinical studies in the United States, the expectations for what information needs to be provided are outlined in 21CRF 312.23(a)(7)(i) and guidance documents discussing the requirements in the different investigational new drug (IND) phases.[5–7] For studies performed in the European Union, similar guidance is provided for Investigational Medicinal Product Dossier (IMPD) content.[8] All the documents emphasize the graded nature of chemistry, manufacturing, and control information needed as development progresses. At all phases, sufficient information should be submitted to ensure the proper identification, strength, quality, purity, and potency of the investigational candidate. The amount of information that will provide assurance will vary with the phases of the investigational compound, the dosage form, the proposed duration of any planned clinical trials, and the amount of information already available about the compound. In addition, developing a defendable strategy for method development and validation can minimize the required resources while providing the appropriate assurance of quality that differentiates the most productive research and development (R&D) analytical group from the others.

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00021-2

Prior to developing any analytical method for finished products, the drug substance must be well characterized and understood. At these earliest stages of drug development, it is essential to gather information on the drug substance such as pK_a, crystal form, melting point, particle size distribution, solubility in different solvents, sink condition and intrinsic drug dissolution rate (useful later for dissolution method development), hygroscopicity, and UV/Vis characteristics. Another critical part of this early development work will focus on forced degradation studies. The results of these studies will help in both formulation development and method development. The first analytical methods that are developed may have minimal validation but should be sufficient to provide analytical support during early formulation development such as excipient compatibility study and prototype formulation development. The results from these early feasibility studies will determine which of the multiple formulation options are selected for further development and, ultimately, clinical study. Therefore, the methods should be able to indicate the significant decreases in potency or increases in related substances. These measures will be helping the selection of the preferred formulations.

In early nonclinical and clinical phases, it is important that the impurity profile observed from each drug substance and drug product lot be established for future reference and comparison. If the impurities present in the nonclinical studies can be confirmed, the information may be useful at later stages in the development course. To be able to generate an impurity profile in a reliable and repeatable fashion, emphasis must be placed on the methods for related substances, as well as the assay method that is used for product release and stability studies. In addition to the need to refer back to the impurity profiles, the material of First in Human (FIH) enabling Good Laboratory Practices (GLP) tox lot is normally manufactured and characterized in the early stages of development. One key component for calculating purity is to have a specific test that separates all impurities. An additional key component of this development is to establish the stability-indicating capabilities of the methods for drug substances and the selected formulation. Characterization of the impurities can be conducted at later stages of development, while in the early stages, identification reliance on the presence of impurities in the FIH enabling the GLP toxicology lot ensures patient safety.

By the time that a project enters Phase I study, sufficient method history for the drug substance and drug product should be available in order to have reasonable confidence in the methodology. At this point, if a method has not yet been validated, formal validation studies should be conducted. For Phase I, methods should be appropriately validated prior to testing the clinical drug substance and the clinical formulation. Partial validation is normally sufficient and acceptable for the early stages. Often, all parameters apart from intermediate precision and robustness are validated at Phase I. Intermediate precision is then completed at Phase II, and robustness is added at Phase III. One validation parameter which is not necessarily required but is strongly recommended is solution stability. By validating a minimum time for which the sample and standard solutions are stable, flexibility in laboratory procedures and the ability to store prepared solutions are achieved. Without validated solution stability, solutions must be prepared fresh prior to each analysis. Additionally, if significant degradation is observed during the solution stability validation experiments, it may indicate possible degradation during analysis.

Another reason for a strategic approach to method validation is that the validation studies may need to be repeated multiple times at different development stages. This revalidation may be required due to new impurities, changes in formulations, or possible improvements in the methodology to utilize new technologies or make methods more appropriate for quality control (QC) laboratories. The full validation of analytical methods to meet the applicable guidance documents is required by the time of filing of application for marketing authorization via the submission of common technical documents.[9] Additionally, every effort should be made to have the final methods in place by the start of the registration stability studies in order to avoid the need to discuss and justify method changes during the submission and review cycle.

The strategic approaches for method validation in each phase of drug development are outlined in Table 21.1.

TABLE 21.1 Phasing Approach for Analytical Method Validation in Each Phase of Clinical Trial Materials

Validation parameters	Preclinical	Phase I	Phase II	Phase III
Specificity	Yes	Yes	Yes	Yes
Repeatability	Yes	Yes	Yes	Yes
Intermediate precision	No	No	Yes	Yes
Linearity	Some	Yes	Yes	Yes
Accuracy	Some	Yes	Yes	Yes
Detection Limit (DL)	Yes	Yes	Yes	Yes
Quantitation Limit (QL)	No	Yes	Yes	Yes
Robustness	No	No	No	Yes
Analytical solution stability	No	Yes	Yes	Yes

21.2 CATEGORY OF ANALYTICAL METHOD AND METHOD DEVELOPMENT

The reliable analytical method is a key element of the quality system. Methods are developed and applied for identification, strength determination, purity measurement, and control of other quality attributes of the products. For the majority of solid oral dosage form drug products, the following tests are required: physical appearance, identification, potency assay, related substance assay, drug release (dissolution) testing, residual solvents (if applicable), and moisture determination. Microbial testing may be required for QC of solid oral products. (Method development and validation for microbial tests is beyond the scope of this chapter.) In addition, other tests may also be required, especially when the API has multiple polymorphs that may have different physical characteristics. These tests could include powder X-ray diffraction (PXRD), near-infrared (IR), and Raman spectrometry. It is extremely important that the polymorphic form of APIs in products is under control during the manufacturing process and stability studies. Additionally, chiral determination may be required during the development.

There are four categories of analytical methods, based on the United States Pharmacopoeia (USP), from analytical determinations to subjective evaluation of quality attributes[10]:

- Category I: Analytical methods for the quantitation of major components of bulk drug substances or active ingredients (including APIs, preservatives) in finished pharmaceutical products
- Category II: Analytical methods for the determination of impurities in bulk drug substances or degradation compounds in finished pharmaceutical products, including quantitative assays and limit tests
- Category III: Analytical methods for determination of performance characteristics (eg, dissolution or drug release)
- Category IV: Identification tests

The data elements normally required for each of the categories of assays will be discussed in Section 21.3. The specific requirement for each individual testing procedure should be taken into consideration and are discussed in the next sections.[2,10]

21.2.1 Identification

As one of the key elements in ensuring the quality of the products, the method intended for identification should be specific for the actives in the drug product. A specific identification test method, such as Fourier transform infrared spectroscopy (FT-IR), is preferred over a chromatographic method. FT-IR is often applied to drug substances. For drug products, the extraction of the actives and subsequent cleanup may be necessary to meet the specifications due to interference from excipients. If nonspecific methods are used, two independent analytical procedures are required. A single nonspecific method, such as UV-Vis or chromatographic retention time using thin layer chromatography (TLC) or high-performance liquid chromatography (HPLC), is not sufficient by itself for identification. However, HPLC equipped with a specific detector like a diode array detector (DAD) or mass spectrometer (MS) becomes a powerful analytical tool, as it obtains both the retention time and UV-Vis spectra simultaneously. Hence, it is considered an acceptable identification technique.

21.2.2 Potency assay

A specific, stability-indicating procedure should be developed and validated. A fully validated procedure, including interlaboratory precision and specificity based on stressed studies, should be part of the phase III method. Reduced validation at earlier phases is applicable, as discussed previously. Generally, it is considered to be acceptable and cost effective for two HPLC assay procedures for commercial products to be developed and validated: one with shorter run time for initial release, and the other with stability-indicating characteristics for stability research. For a method of initial release, it is required to ensure that the active peak is well clear of all known impurities, including those that may arise from the manufacturing process, while other degradants are unlikely to be present in the fresh products.

21.2.3 Impurities

Impurities in drug substances and drug products include organic impurities, inorganic impurities, and residual solvents. Organic impurities are generally called *related substances* since they are structurally related to the drug substance. These substances may be identified or unidentified degradation products and can arise from many sources, including the raw materials, starting materials, and purchased intermediates, during the manufacturing process or storage of a material. The earlier the impurities are identified during drug development, the more time there is to address them through process changes or qualification studies. Guideline Q3B outlines the identification and qualification thresholds for impurities in drug

substances and drug products. During drug development, validation of impurity methods for accuracy, detection limits (DLs), and quantitation limits (QLs) and response factors can be accomplished by spiking the impurities into the drug product or into a placebo mixture and by evaluating the analytical results of the spiked samples. In addition to the majority of impurities, see ICH guidance M7, which covers assessment and control of potential genotoxical impurities.[10–13]

21.2.4 Dissolution

Dissolution or drug release testing is commonly considered to be a critical QC for solid oral dosage forms. In the absence of in vivo data, it is a challenge for scientists to develop a procedure that not only guides the formulation development process but also can be used as a regulatory test to detect manufacturing deviation and ensure product consistency both upon release and over the products' shelf life. The primary goal is to develop a discriminating, rugged, and reproducible dissolution method, which must be able to highlight significant changes in product performance due to changes in the formulation or manufacturing process. The objective of dissolution testing varies during the life cycle of dosage form development. During the early stage of phase I, a method is developed to establish the mechanism of in vitro drug release. During phases II and III, the objective shifts to identifying a test method that can provide an in vitro–in vivo correlation (IVIVC) or other biorelevant information. By having a method that has a proven IVIVC, a significant increase in the guiding power of the dissolution method for developing bioequivalent multiple strengths of the product and a reduction in the regulatory burden caused by postapproval changes can be achieved. The inclusion of a new chapter in USP 30/NF 25 provides recommendations on how to develop and validate a dissolution procedure.[22] See chapter "Dissolution Testing of Solid Products" for more information on dissolution testing of drug products.[14–25]

21.2.5 Blend homogeneity and dosage uniformity

The term *uniformity of dosage unit* is defined as the degree of uniformity in the amount of the drug substance among dosage units. The harmonized compendial chapter on uniformity of dosage units outlines the release criteria to show uniformity. In addition to the uniformity at release, it is important during development to demonstrate the uniformity of blend samples and the uniformity of dosage units throughout the manufacturing run (lot uniformity). The analytical procedure for assay can be used to determine the drug content of uniformity samples. In the cases where a complicated procedure or long HPLC run time is required for assay but is not needed for blend uniformity/content uniformity (BU/CU) determinations, a separate testing procedure with shortened run time may be developed. Run time in the BU/CU chromatographic methods may be much shorter than in the assay method because stability-indicating capability may not be required. Generating BU and lot uniformity data is the key to understanding the manufacturing process during the development phases of the products. The procedures to demonstrate adequacy of mixing for powder blends using stratified sampling of blend and dosage units were published by Boehm et al. (2003)[26].

21.2.6 Cleaning test method development

Regardless of the stage of development, if reusable equipment is used during manufacturing, current good manufacturing practices (cGMP) require that cleanliness of the equipment be demonstrated to eliminate the risk of cross-contamination and to ensure that residual material is removed from the equipment after use. The validation of analytical procedures used to determine residual actives from swabbing or rinsing samples should be conducted as a critical step for cleaning validation.

In addition to the normal validation required for analytical methods, cleaning analytical methods must include confirmation of surface recovery. Depending on the equipment and how cleaning confirmation is performed, surface recovery studies may be conducted using both rinsing and swabbing samples from multiple surface types. For many types of equipment, cleaning is performed using large volumes of the appropriate solvent. A portion of the final rinse solvent can be collected and taken for analysis. In this case, confirming recovery based on rinsing is appropriate. For some pieces of equipment, the surface is swabbed and the swabs are taken for analysis following rinsing. In these cases, surface recovery using swabbing is appropriate. Common manufacturing materials include different grades of stainless steel (some of which have many different recovery characteristics), polymers, and glass. Spiking and recovery experiments need to be performed on coupons of the materials that will be used in manufacturing.

There are a number of challenges in validating a cleaning method. To reduce the possible challenges, scientists in analytical labs should work closely with

the manufacturing group to understand what is required for the method and eliminate unnecessary hurdles. These hurdles could include the use of solutions like bleach, which destroys the active molecule but results in degradation products that could have toxic characteristics; cleaning solutions that interfere with the chromatography or degrade the active substance in solution; or swabs that have extractables that interfere with the analysis. To reduce the risk of extra unknown peaks from the materials of swabs, the same materials that will be used in the manufacturing facility must be used during method development. The selection of the solvent should be based on the solubility of its active ingredient and ICH classification. The first choice of solvent should be water, isopropyl alcohol (IPA), or a mixture of the two. Minimum recovery requirements can be debated, but recoveries of <50% are difficult to justify. The final analytical method should include an adjustment factor based on the worst-case recovery to ensure that reported results are a true indication of the amount of residual drug substance left on the surface.

If water and IPA are not viable cleaning solvents, solvents that better solubilize the drug substance or cleaning agent may be required. Class 2 solvents should be avoided, and Class 1 solvents should never be used. The proposed solvents, which are commonly used, are listed in Table 21.2[12]. Depending on the final cleaning solution selected, it may be necessary to add an additional cleaning step with IPA, water, or both to the areas that were swabbed to ensure that all other solvents are appropriately removed.

TABLE 21.2 Proposed Swabbing Solvents

Number	Solvent	Boiling point (°C)	Relative evaporation rate[a,b]	Remarks[c]
1	Isopropanol	82.6	1.7	Class 3
2	Ethanol	78.1	1.95	Class 3
3	0.25% Aqueous acetic acid			Pharmaceutical acceptable excipient
4	Acetone	56.2	6.06	Class 3
5	Ethyl acetate	77.3	4.94	Class 3
6	Isopropyl acetate	85	3.0	Class 3
7	Methanol	64.6	5.91	Class 2
8	Acetonitrile	81.7	2.33	Class 2

[a] *All BPt and relative evaporation rate values were taken from MSDSs.*
[b] *All relative evaporation rates are relative to* n*-butyl acetate.*
[c] *Classification as per ICH Q3C for residual solvents.*

21.2.7 Other analytical techniques

Thus far in this chapter, we have focused on the key tests required to ensure the key quality attributes of solid oral dosage forms. In addition to the tests described up to now, many other tests may be required during development, and even as part of the filing to ensure the quality of the drug product. All methods should be appropriately validated for their intended use, and the amount of validation depends on the type of method and the stage of development. Some of the other methods that may be used include (but are not limited to) near-IR, Raman spectrophotometry, PXRD, gas chromatography (GC), optical rotation, and chiral chromatography; all of these could be applied to specific testing as well.[27–33] For example, PXRD may be used to monitor the change of polymorphs of drug substances in drug products during manufacturing process and storage.

21.3 ANALYTICAL METHOD VALIDATION

The analytical method must be validated for the intended use of the substance and will be applied for the release of clinical trial materials, market products, or both. Once the analytical method is validated and approved, it will be used to monitor the quality of a product over its life cycle. However, the performance of the method should be evaluated for appropriateness from time to time. The method may be required to be optimized by adjusting and revalidating the operational conditions. New methods may be developed and validated using new technologies (due to cost effectiveness and sensitivity) or required due to awareness of a new impurity or a recent decrease in the impurity limit. Therefore, analytical method development and validation are considered part of the lifecycle management of a product.

21.3.1 Verification of compendial methods

A compendial procedure is considered validated if it is published as official text in a pharmacopoeia such as the USP-NF, in a supplement, or as an interim revision announcement in *Pharmacopeial Forum (PF)*. When using compendial methods, full validation is not necessary, but verification of the procedure is very important. Verifications ensure that the procedure is suitable for

use with a specific ingredient or product, in a specific laboratory.[34] For example, titrimetric methods for water determination should be verified for accuracy (and absence of possible interference) when used for a new product or raw material. For impurity testing, the suitability of a compendial procedure may be an issue for several reasons (eg, impurity profile changes from different routes of synthesis, composition of formulation, or interference from excipients). It is recommended that the testing lab thoroughly evaluate the suitability of the pharmacopeia monographs prior to being applied to their own products.[35–37] For example, forced degradation may be applied to establish the suitability of the chromatographic conditions.

21.3.2 Characterization of reference standard

During method validation, a well-characterized standard should be used. A well-characterized reference standard is a critical factor for method validation. For potency assays, the purity of the standard must be assigned. For impurity assay validation, some impurity standards may not be of high purity, but they must be characterized to confirm their identity. It is important to ensure that any impurities contained in the impurity standard do not interfere with the analysis. If possible, the reference standard (used as the primary standard) should always be obtained from a recognized authority, such as the National Institute for Standard and Technology (NIST), USP, or European Pharmacopeia (Ph.Eur.). For new drugs, this is rarely possible; therefore, an in-house reference standard must be synthesized and characterized. The reference standard should be characterized by at least the following tests: physical appearance, identification, and purity assignment. The structure should be confirmed using multiple analytical techniques such as elemental analysis, IR spectroscopic analysis, UV-vis spectroscopic analysis, mass spectroscopy, ^{1}H-NMR, and ^{13}C-NMR. Purity assignment can be established by testing for the following: organic impurity, inorganic impurity, moisture, and residual solvents. The amount of total organic impurities may be determined by HPLC or other chromatographic methods. Moisture and residual solvents may be determined by thermogravimetric analysis or a combined Karl Fischer titration and GC method. The amount of inorganic residue can be determined using residue upon ignition. If the Karl Fischer method for the determination of water content and the GC method for residual solvents are applied, these methods should be validated (or verified if using a compendial method) prior to use. The in-house standard can be qualified against a primary standard following a well-defined qualification protocol.

21.3.3 Stability-indicating method

A stability-indicating assay, as required per the ICH guidelines,[38] should be used for any GMP stability program so that the procedure is able to detect the changes in the pertinent properties of the drug substance and drug product over time, without interference from degradation products, process impurities, excipients, or other potential system components. It is important that a comprehensive forced degradation study and HPLC coelution evaluation be conducted in order to demonstrate the ability of the procedure to detect any changes that are attributable to degradation.

21.3.4 HPLC coelution peak evaluation

HPLC with DAD or MS is a powerful tool that has the ability to detect very low levels of coeluting components. These techniques have the ability to evaluate the purity/integrity of chromatographic peaks via the HPLC method, which can be used in development and validation and are widely used in the pharmaceutical industry. Since impurities in the APIs and the actives themselves often have very similar chemical structures, the impurities are also likely to have very similar chromophores and ultraviolet (UV) spectra as the APIs. For this reason, the software provided with DAD generally includes algorithms that calculate peak purity. Additionally, during method development, forced degradation studies are performed to generate samples with exaggerated levels of likely impurities. These samples are used to evaluate and identify when coeluting components are present, at which time the increased level will offer the algorithm a greater chance of success. It is important to note that the peak purity check can only prove that the peak is impure or contains a coeluting component. DAD cannot prove that the peak is absolutely pure. However, if the peak purity check passes, the peak purity is normally considered acceptable.[39]

21.3.5 Forced degradation studies (stress studies)

Forced degradation studies, or stress studies, are the main tools that can be used to predict stability issues, develop analytical methods, and identify degradation products and pathways.[40] These studies should be performed prior to other validation parameters (eg, accuracy, repeatability, intermediate precision, specificity, DL, QL, linearity and range, solution stability, and robustness). An approved protocol is recommended before forced degradation studies can be performed.

For multiple strengths of drug products with the same excipient composition (including different ratios), forced degradation studies can be performed with

only one formulation. For multiple strengths or a formulation of drug product with different excipients, each different formulation composition should be evaluated using forced degradation studies.

If the assay methods for drug substances and drug products have different conditions that can cause changes in selectivity (eg, different HPLC columns, mobile phase compositions, gradient programs, flow rates, column temperatures, etc.), forced degradation studies should be done for both drug substance and drug product methods. In early IND phase products, if the assay methods for the drug substance and drug product have the same conditions, which result in the same selectivity, it is preferred to do forced degradation studies with drug substances for IND Phase 1 products, and with drug products for IND phase 2 products.

For IND Phase 3 phase applications and later, forced degradation studies should be done for both drug substances and drug products. Additionally, mass balance should be evaluated by the related substances detected to the assay results obtained for forced degradation samples. Generally, 5−20% degradation is optimal for forced degradation studies, although an excessive degradation is acceptable so long as it does not cause the peak purity check to fail. The more a sample is degraded, the more likely a loss in mass balance due to secondary degradation, loss of impurities in the solvent front, and loss of absorption due to ring-opening or other degradation pathways will be observed. The forced degradation studies for drug product should be performed before commencing stability studies of registration batches. The stress conditions described next are considered as extreme conditions. Alternative or less stressful conditions may be applied when excessive degradation or interference is observed.

During data acquisition, DAD should be used for HPLC, and all spectra of peaks should be collected.

The samples should include the drug substance, placebo, and drug product, and may be subject to the following conditions:

Acid/base:	Drug substance as is; 0.1 N of HCl/NaOH at RT for 1 day
Thermal:	Composite sample; 85°C for 10 days
Thermal-humidity:	Composite sample with 0.5 mL of D.I. water in a crimped vial; 85°C for 10 days
Photostability:	One layer of composite sample (thickness ≤3 mm; use a petri dish with a quartz plate cover, an open petri dish, or a dish covered with parafilm), 2−3× ICH option 2.[41]
Oxidation:	Composite sample; 3% H_2O_2 at RT for 10 days, light-protected

The following results must be included in the report: (1) assay results for parent compounds, (2) the results of coelution test of parent compounds, and (3) the profile of degradation products and other related compounds.

For Phase 3 studies, the mass balance results should also be described.

21.3.6 Method validation parameters for chromatographic methods

The steps for method validation are as follows:

- The protocols or general standard operating procedures to describe the detailed parameters, as per Table 21.3
- Execution of the protocol in the lab
- Redevelopment and validation if deviation or failure is observed during validation that is attributed to the analytical method
- Validation report

Validation parameters are included (but not limited to) linearity, precision, accuracy, robustness, QL, DL, and stability of the analytical solution. The analytical parameters such as system suitability and filter bias should be evaluated by the chromatographic method during method validation.[2,42]

21.3.6.1 Filter bias

In modern analytical techniques, the analysis is typically carried out by spectroscopy (UV-vis) utilizing a chromatographic method, either HPLC or ultra-performance liquid chromatography (UPLC). Often a filter or centrifuge is utilized to remove particulates that may clog the column or affect absorbance readings. The different types of syringe filters such as those made of nylon or polytetrafluoroethylene (PTFE), with a size ranging from 0.45 to 0.2 μm should be investigated depending on the samples. The filter should be validated by filtering a portion of the working standard solution or sample solution through each syringe filter, discarding the first 2−3 mL, and collecting the filtrate for analysis. The result from the filtered solution should be comparable to that of the unfiltered solution.

21.3.6.2 System suitability

System suitability demonstrates that the system is working properly at the time of analysis. The appropriate system suitability criteria are based on individual technology and the samples being analyzed. For all chromatographic procedures, system suitability should include injection repeatability expressed as the relative standard deviation (RSD) of peak responses obtained from five or six consecutive injections of a working

TABLE 21.3 Data Elements Required for Analytical Method Validation

Validation parameters	Identification method	Related substances and residual solvent methods		Assay, dissolution, and preservative methods
		Quantitation	Limit	
Specificity	Yes	Yes	Yes	Yes
Linearity	No	Yes	No	Yes
Range	No	Yes	No	Yes
Accuracy	No	Yes	No	Yes
Precision	No			
Repeatability		Yes	No	Yes
Intermediate precision		Yes[a]	No	Yes[a]
Detection limit	No	No	Yes	No[b]
Quantitation limit	No	Yes	No	No[b]
Robustness	No	Yes	No	Yes

[a]*Intermediate precision may be omitted if reproducibility is performed during method validation.*
[b]*Detection limit or quantitation limit may be required for methods such as determination of absence of active in placebo, dissolution testing of modified release drug products, eg, acid resistance of enteric coated tablets.*
Note: The system suitability should be evaluated for all quantitation methods.

standard solution, tailing factor (T), theoretical plate number, resolution, and relative retention time (RRT).[42] For nonchromatographic methods, system suitability tests are also used to confirm that the system is functioning correctly. For example, titration should always include a blank titration. For some particle size analyses, the solution with a known particle size is used to ensure that the system is functioning as expected.

Since HPLC methods are the most commonly used procedures, the parameters for system suitability are injection repeatability, check standard, T, theoretical plate number, system drift, and resolution, as discussed in detail in the following sections. Other system suitability parameters may be required, such as the peak-to valley ratio and signal/noise ratio for impurity testing.

21.3.6.2.1 Injection repeatability

The working standard solution will be injected five or six times into the chromatographic column. The mean and RSD for concerned peak responses such as area will be calculated as follows:

$$X(mean) = \sum_{i=1}^{n} X_i / n$$

$$\%\text{RSD} = \frac{100}{X(mean)} \sqrt{\frac{(X_i - X)^2}{n-1}}$$

where

X_i = peak response of individual injection
n = number of the repeatable injections

In general, the acceptance criteria are as follows:

The RSD for the peak area of interest from five or six injections of the working standard solution should be $\leq$2.0% for potency assay, $\leq$10% for impurity testing and residual cleaning testing, and $\leq$3.0% dissolution testing.
If the product has low strength or S/N of the active peak is less than 50, the RSD of the peak area of the active from the six consecutive injections of $\leq$3% may be acceptable for potency assay.
For API potency testing, a maximum permitted RSD is defined in USP based on specification limits.

21.3.6.2.2 Check standard

A check standard is often considered as part of integrated system suitability for a potency assay or dissolution assay. The percent recovery of the active from the check standard is calculated as follows:

$$\%\ \text{Recovery} = \frac{A_{CK}}{Wt_{CK}} \times \frac{Wt_{WSTD}}{A_{WSTD}} \times 100$$

where

A_{CK} = the peak area of the active from the check standard solution
A_{WSTD} = the average peak area of the active from the six injections of the working standard solution

Wt_{CK} = weight of the active used in the preparation of the check standard solution (mg)
Wt_{WSTD} = weight of the active used in the preparation of the working standard solution (mg)

It is acceptable that the percent recovery for the check standard solution is in the range of 100.0 ± 2.0. For a low dose/strength of the drug product, a 3.0% difference may be acceptable.

Additional standard solutions may be needed in some cases. For example, a QL or DL solution for an impurity test should be injected to ensure that the method is sensitive enough to detect impurities at the required concentration. The sensitivity could vary due to different laboratories, temperatures, columns and detectors. For a DL solution, a peak must be detected and integrated. For QL solutions, the peak area should be compared to the standard peak for recovery. For a QL solution, the recovery should be between 80.0% and 120.0%.

21.3.6.2.3 Tailing factor

The T of the active peak from the working standard solution is calculated as follows:

$$T = \frac{W_{0.05}}{2f}$$

where $W_{0.05}$ is the peak width at 5% of the active peak height from the baseline, and f is the distance from the peak maximum to the leading edge of the peak (the distance being measured at a point 5% of the peak height from the baseline). The peak is symmetrical if $T = 1.0$. It is generally considered to be reasonable if T is no more than 2.0, but the acceptable T should be confirmed during robustness experiments since excessive tailing or fronting could affect resolution between peaks.

21.3.6.2.4 Theoretical plate number

The theoretical plate number per column (N) for the peak can be calculated from the first injection of the working standard solution as follows:

$$N = 16\left(\frac{t}{w}\right)^2$$

where t is the retention time of the active peak and w is the peak width of the peak, which is obtained by extrapolating the relatively straight sides of the peak to the baseline. The appropriate requirements for this parameter should be derived from validation data.

21.3.6.2.5 System drift

Periodic injections of the working standard solution should be made after a certain number of sample injections and at the end of the run. The percent recovery of system drift injection can be calculated as follows:

$$\%\ \text{Recovery} = \frac{A_{SCK}}{A_{WSTD}} \times 100$$

where

A_{SCK}: peak area of the active from the system drifts injection of the working standard solution
A_{WSTD}: average peak area of the active from the first six system suitability injections of the working standard solution

For system drift, the percent recovery of system drift injection throughout the run should be the same level as that of the injection repeatability. If the system drift meets this requirement, the average peak response from the first consecutive injections can be used to calculate the samples. Otherwise, a bracketing procedure should be used. In addition to the peak area, the retention time should be evaluated. For identification, the retention times should not vary by more than 2%.

21.3.6.2.6 Resolution

The resolution factor (R) should be calculated to demonstrate that the critical pairs of adjacent peaks are adequately separated and therefore can be independently integrated. The value of R from the first injection of the working standard solution can be calculated as follows:

$$R = \frac{2(t_{n+1} - t_n)}{W_{n+1} + W_n}$$

where

t_n = retention time of peak n
t_{n+1} = retention time of peak $n + 1$
W_n = peak width of peak n
W_{n+1} = peak width of peak $n + 1$

The minimum resolution between each identified critical pair of adjacent peaks should be ≥ 1.5 since this represents the baseline separation of two neighboring Gaussian peaks. Resolution can also have an upper limit to make sure that peaks have not drifted too far away, minimizing the risk of significant changes in the retention properties of a column.

21.3.6.3 Accuracy

The accuracy of the method can be determined by spiking known amounts of the active at suitable levels of the label-claimed amount, as per Table 21.4, to the corresponding placebo powder (or an amount of placebo mixture containing all the ingredients for the

TABLE 21.4 Accuracy and Acceptance Criteria Proposed

Test	Accuracy	Acceptance criteria
Cleaning residual	70%, 100%, and 130%	Minimum lowest recovery at each level ≥50%
Potency (DS and DP)	70%, 100%, and 130%	Average recovery at each level is between 98.0 and 102.0
Dissolution (immediate release)	50%, 100%, and 120%	Average recovery at each level is between 97.0 and 103.0
Dissolution (controlled release)	20%, 100%, and 120%	Average recovery at each level is between 97.0 and 103.0
Dissolution (delayed release)	Acid stage: 1% and 10% Buffer stage: 50% 100%, and 120%	Average recovery at each level is between 97.0 and 103.0
Related substance (DS and DP)	QL, 100% and 200% of spec.	Average recovery at each level is between 90.0 and 110.0 For QL, 80.0−120
Residual solvent	50%, 100%, and 120% of spec.	Average recovery at each level is between 80.0 and 120.0

formulation except active), and then calculating the percent recovery of the active. For Phase 3 accuracy experiments, three sample preparations are required at each spiking level, and a minimum of three levels should be assessed. Table 21.4 describes the range and the levels for each individual testing procedure.

Percent recovery is calculated by the assayed amount divided by the known amount of analyte spiked in the sample. Typical acceptance criteria are also provided in Table 21.4. Percent recovery, average recovery from each level and overall levels, and confidence intervals should be evaluated.[43] For potency assay, the RSD of recoveries for each spiked level should be not more than 2.0%. For low strength of drug products (eg, ≤1 mg), a wider range such as 3.0% for potency and 5.0% for dissolution may be applied. For dissolution testing, the drug substance and placebo (or an analytically prepared placebo) may be spiked into the dissolution vessel. The mixture of drug substance and placebo may be considered to be useful if the recovery of the analyte of interest is less than expected. It could be due to the unique characteristic of drug substances in the dissolution media. For low strength of the drug product, it may not be possible to spike drug substance into the dissolution vessel, in which case spiking with drug solution may be used. Drug substance solutions with high concentrations should be prepared by dissolving in an organic solvent or dissolution medium. The dissolution medium to use is the first choice. If the medium is not suitable, organic solvents may be used, but the volume of solvent spiked should not be more than 5% of the total medium volume.

For recovery studies in impurity tests, an aging API or stressed API sample spiked into placebo may be used to demonstrate the recovery of impurities in the sample when the impurity standards are not available.

21.3.6.4 Precision

Precision should be evaluated through repeatability, intermediate precision, or reproducibility. These concepts are discussed in the next sections.

21.3.6.4.1 Repeatability

For related substances and residual solvent tests, repeatability can be evaluated by spiking impurities or solvents at the specification limit for the products. In early phases of development, little may be known about the impurities, so minimal repeatability data may be generated. For the potency assay, repeatability can be assessed using drug products (tablet or capsules) with a minimum of six determinations at 100% of the test concentration. Typically, the RSD of 2.0% for the % label claim (LC) of the active from six replicate sample preparations is achievable.

21.3.6.4.2 Intermediate precision

Intermediate precision is a measure of the method's sensitivity to minor changes in equipment performance and to variations in the operator's technique on a given day. A second analyst should perform the assay using different equipment and on a different day to confirm that acceptable results can be obtained in the intralaboratory settings. The absolute difference between the mean % LCs of the active generated by the variation of these factors, including analysts, time, and equipment, should be <3.0%, but the exact criteria are based on the type of test and ultimate specification (eg, if the drug substance assay specification is 98.0−102.0%, then the difference between laboratories should be <1.5%, but for a drug product with a specification of 90.0−110.0%, a wider criteria such as NMT 3.0% could be used).

21.3.6.4.3 Reproducibility

Reproducibility is a measure of the method's sensitivity to laboratory changes. It could be moderate changes in equipment performance or variation in the operator's technique and the lab environment. Reproducibility is generated by two separate laboratories running the test and is therefore also called *interlaboratory precision*. The absolute difference between the

mean % LCs of the active generated by the two labs should be <3.0%, but the exact criteria are based on the type of test and ultimate specification (eg, if the drug substance assay specification is 98.0−102.0%, then the difference between laboratories should be <1.5%, but for a drug product with a specification of 90.0−110.0%, a wider criteria such as NMT 3.0% could be used).

The results, such as mean, standard deviation, RSD, and confidence interval, should be evaluated and reported on for each type of precision.

21.3.6.5 Linearity

A linear relationship between the concentration and the respective response can be obtained by analyzing a series of standard solutions.

At least five standard solutions with a specific range, as specified in Table 21.5, should be prepared. One injection of each of the linearity standard solutions may be sufficient. The peak area of the active will be measured at different concentrations and plotted against the corresponding concentrations. The correlation coefficient (r), y-intercept, slope of the regression line, and residual sum of squares will be calculated by the method of least squares, and a plot of the data should be reported.

The correlation coefficient should not be less than 0.999 for potency assay, and not less than 0.99 for other tests. It is very useful to evaluate the difference between the estimated value from regression line and the actual value. Limits for the difference may be defined when applicable. For example, it is acceptable that the percent difference between the calculated value (concentration) from the regression line and the actual value is no more than 20% at the QL level for impurity assay.

In some cases, a combined method for potency and impurity assay is used. It is quite common that impurity standards are unavailable in the early stage of product development. The area percent or amount of impurity calculated from the standard at 100% level of LC is typically reported if it is linear from the reporting or the QL level to 120% of the label-claimed level of strength. Otherwise, underestimated results may cause some issues later if the degradation/impurity approaches or exceeds the qualification/identification level as defined in ICH Q3B(R2). If it is not linear from the QL to 120% of LC, a diluted standard should be used to minimize any bias in the impurity calculation. If impurities are available, impurity standards containing known impurities at the respective specification levels should be prepared. By comparing the slope of each impurity to the slope of the main standard, response factors or normalization factors can be established and a single point standard can be used to accurately quantitate the known impurities.

TABLE 21.5 Ranges in the Regression Line

Test	Linearity
Cleaning residual	QL to 130%[a]
Potency (DS and DP)	70% to 130% of LC
Dissolution (immediate release)	50% to 120% of LC
Dissolution (controlled release)	20%[b] to 120% of LC
Dissolution (delayed release)	Acid stage: QL to 10%; Buffer stage: 50% to 120%
Related substance (DS and DP)	QL or reporting level to 200% of spec.[c] QL or reporting level to 120% of LC if area percent is reported.
Residual solvent	50% to 120%[d] of spec.

[a]*Consider 200% when spec. is very low.*
[b]*Unless the spec. requires lower %*
[c]*If spec. is unknown, use 0.5%.*
[d]*Consider 200% when spec. is very low.*

21.3.6.6 Specificity

The specificity of the method must be investigated in order to confirm that an analytical procedure is specific for the analyte of interest in the presence of components such as impurities, degradants, and matrix components (excipients). For an HPLC procedure, the specificity can be demonstrated by the separation of the critical pair of components (i.e., API and impurity, two impurities, etc.). In such cases, a DAD is useful for detecting coeluted peaks in the samples spiked with an impurity when the impurity is available, as well as in the stressed samples. Also, a carryover study should be conducted. For HPLC methods, it is recommended to double the isocratic run time or the holding time of the last stage of a gradient elution program to detect any strongly retained eluents in the placebo or samples.

21.3.6.7 Stability of standard and sample solutions

It is essential that the sample and standard solutions are stable throughout sample preparation and analysis. Proving the stability of the standards should be part of the validation process.

The standard solutions should be freshly prepared, and the concentration of the standard solution is used as the initial value. Portions of the working standard solutions are stored under refrigeration (5°C ± 3°C) and at controlled and monitored room temperature.

These stored portions of the standard are assayed at various time points (such as 1 day, 3 days, etc.) to determine the concentration of the analyte of interest with reference to the freshly prepared standard solution.

For potency assays, the standard and sample solutions are considered to be stable if the percent difference between the initial values and those at a specific time is no more than 2.0%, but any downward trend in the data should also be evaluated for possible impact on the analysis.

For dissolution samples, the stability of the samples in the dissolution vessels is another parameter to be validated in addition to the stability of the standard and sample solutions discussed previously. The sample in the dissolution vessel should be stable at least up to the final sampling time. The sample solution can be obtained by spiking the analyte of interest (either by spiking a drug solution into the medium or by dissolving the API substance into the medium in the vessel within a very short time) and placebo at 37.0°C ± 0.5°C.

For impurity analysis, the sample solution is considered stable if the following conditions are met:

If 0.10% ≤ the individual related substance <0.50%, the absolute difference between the initial and specific time point values should be ≤0.10%.
If % individual related substance ≥0.50%, the percent difference between the initial and *t*-time point values should be ≤20.0%.
No new degradation product ≥ QL of analyte of interest should be detected.

If these conditions cannot be met at any of the time points and storage conditions, then the sample and standard solution must be analyzed within the time period in which these conditions do apply. In the worst-case scenario, the sample solution has to be freshly prepared prior to each injection to obtain consistent impurity profile results.

21.3.6.8 DL and QL

As defined in ICH Q2 (R1), the DL of an individual analytical procedure is the lowest amount of analyte in a sample that can be detected, but not necessarily quantitated as an exact value. The QL is the lowest amount of analyte in a sample that can be quantitatively determined with suitable precision and accuracy. The DL and QL are critical parameters of analytical procedure validation for residual solvent and impurity assay.

There are several approaches for determining DL and QL, which are described in ICH Q2 (R1) and discussed in the next sections.

21.3.6.8.1 Visual evaluation

Visual evaluation is most likely to be used for the noninstrumental method. They can be determined by the analysis of samples with known amounts of an analyte and by establishing the minimum level at which the analyte can be detected (DL) or quantified with acceptable accuracy and precision (QL). For QL determination, six replicate samples may be required to be prepared and tested.

21.3.6.8.2 Signal-to-noise-ratio approach

The signal-to-noise ratio of the peak of analyte of interest in the sample should be at least 3:1 from DL solution and 10:1 from the QL solution. For chromatographic techniques, the signal of the peak and the baseline noise can be measured manually or instrumentally using the built-in software. It is typical to follow USP procedures for the determination of signal-to-noise ratios.

For HPLC and GC, the DL and QL of an analyte may be determined by the serial dilution of a standard solution with diluent and injecting into the chromatographic systems for assay. Then QL and DL are determined by the signal-to noise ratio.

21.3.6.8.3 Standard deviation of the response and slope approach

DL and *QL* may be expressed as follows:

$$QL = \frac{10\sigma}{S}$$

$$DL = \frac{3.3\sigma}{S}$$

where

σ = the standard deviation of the response
S = the slope of the calibration curve

In this instance, σ may be estimated based on the blank standard deviation (measurement of the magnitude of analytical background response using six replicate blank samples), residual standard deviation of regression line, or the standard deviation of *y*-intercepts.

The reporting levels are defined as per ICH Q3B(R) based on daily intake and dose. The reporting threshold for impurity testing, 0.05%, is considered as the lowest level in most cases except for potential genotoxic impurities. Therefore, it is prudent that the concentration of the analyte in sample solution equivalent to 0.05% is prepared as the QL level. The signal-to-noise ratio for the peak of analyte should be more than 10:1. Then, QL solution should be injected as six replicates onto a chromatographic system for analysis. The RSD of the peak area of the analyte should be

≤20.0%. If the QL level cannot meet any of these criteria, the 0.10% level will be evaluated. If the 0.10% level still cannot meet any of the criteria, the method should be modified to meet all of the QL criteria.

21.3.6.9 Robustness

The robustness of an analytical procedure is critical for effective method validation and cost effectiveness in routine assays performed later. For abbreviated new drug application applications, it is recommended that the robustness of a method be evaluated during early stages of method development in order to better understand method performance over changes in operational parameters and samples. If development data support the validation of a method, they can be submitted in a validation report.

Robustness of chromatographic conditions will be performed on a sample solution, such as a repeatability sample solution with injections in triplicate, by varying the parameters specified in Table 21.6. Only one parameter at a time is altered, while the rest of the parameters remain unchanged. The design of experiment could also be used to allow multiple parameters to be varied in each experiment and thus reduce the number of experiments.

The instrument system must be equilibrated under each target and robustness condition. The system suitability requirements should be evaluated in each experiment to ensure the appropriate system suitability criteria are set for the method. Triplicate injections of the sample solution are then made under each condition. The individually determined mean values and the RSD of three injections for each robustness condition are reported. The percent of target value is calculated for each robustness test condition and should be 98.0−102.0%. If any of these robustness parameters fail to meet the acceptance criteria, a precautionary statement should be included in the method specifying the limitations.

When possible, equivalent research on different columns (lots or suppliers) should be evaluated. For related substance methods, additional evaluation should be made to determine the impact of variations on the specificity of the method.

21.3.6.9.1 Robustness on sample preparation

Robustness of the sample preparation for potency assays will be performed in duplicate sample preparations by varying the parameters specified in Table 21.7. Only one parameter is varied at a time while maintaining the remaining parameters at a target condition.

Robustness of the dissolution method will be performed on three tablets (or capsules) by varying the parameters such as medium pH value (±0.1 of target) and concentration of surfactant (±5% of target).

TABLE 21.6 Experimental Design for Chromatographic Parameters

	Parameters	Robustness test condition
HPLC	Mobile phase composition	Target: ±1%
		eg, Target: 50:50 (v/v0)
		Test-1: 49:51
		Test-2: 51:49
	Buffer pH (mobile phase)	Target: ±0.1
	Buffer molarity (mobile phase)	Target: ±10%
	Column temperature	Target: ±3°C
	Wavelength	Target: ±2 nm
GC	Temperature	±5°C
	Flow rate	10% of target

21.3.7 Nonchromatographic method validation

Additional tests to control the quality of the drug substances, excipients, and drug products such as particle size distribution, optical rotation, and methodologies such as DSC, PXRD, Raman spectroscopy, and near-IR spectroscopy should be validated prior to use. The validation parameters may be less extensive than for chromatographic methods, while in general they are likely to include repeatability, intermediate precision, and robustness; however, specificity and accuracy may also be applicable.

21.3.8 Failure and revalidation

During method validation, failure to meet validation acceptance criteria may be observed. If this occurs, investigations should be conducted to track the root causes. If these are attributed to the procedure, the validation will be terminated at this stage and new method development or revision will be initiated. The method is then appropriately revised and all changes are documented prior to revalidation. In case of a failure of comparison during repeatability and intermediate precision, all aspects should be taken into consideration (ie, chemist skills, lab equipment

TABLE 21.7 Experimental Design for Assay Sample Preparation

Parameters	Test conditions
Shaking time	Target: ±20%
Sonication time	Target: ±20%

variation, and sample variation). For any failures during validation experiments, a thorough evaluation is required to ensure that the failure is truly due to the method and not due to laboratory error or other unexpected issues like sample homogeneity.

The need to revalidate the method will be evaluated if there are changes such as column vendor, drug substance route of synthesis, and drug product composition. Some changes may not require revalidation or may require only partial revalidation (eg, new excipients would require specificity experiments), but the evaluation should always be made and the justification for not revalidating should be documented.

21.3.9 Life cycle management of test procedure

Life cycle management of approved test procedures is recommended in recently published FDA guidance for industry-analytical procedures and method validation for drugs and biologics.[2] As part of the continued improvement of quality of the products, the suitability of existing method performance should be evaluated at regular intervals due to manufacturing process changes, detection of new impurities, and decrease of specification limits of some impurity over a life cycle of the product. New test procedures may be developed and validated due to the emergence of new technologies to reduce cost and enhance efficiency. See the FDA guidance[2] for a detailed discussion of the requirements for life cycle management of analytical procedures regarding revalidation, analytical method comparability studies, and reporting postmarketing changes.

21.4 METHOD TRANSFERS

21.4.1 Definition

Any additional analytical laboratory that was not a part of the original validation process must be qualified to demonstrate that the lab executes the analytical procedure in an equivalent fashion to the originating lab. The method transfer usually occurs between the analytical R&D lab and a QC lab. This transfer can be within the same company but could also be between two different companies or different sites. There are two approaches of method transfer. Interlaboratory qualification (ILQ) is a common approach. The testing was performed on same lot of samples or spiking samples by two labs and the results obtained are comparable. Among other approaches is full or partial method validation by receiving lab. All involved methods must be assessed in order to determine the status of method transfer as necessary. Under certain circumstances, method transfer waiver for receiving lab is eligible (refer to USP38 <1224> for the requirement).

It is quite realistic or practical to avoid method transfer as a separate study if a QC laboratory or receiving lab can participate in the initial validation process or covalidation such as a reproducibility study. Under certain circumstances, it is very important to decide which approach of method transfer is more appropriate. It might be the case that partial validation of the method is more appropriate in early phases of drug product development between a contract research organization (CRO) and a sponsor, and ILQ in later phase stages.

The following steps will be required for method transfer:

1. Analytical procedure and testing materials, including reagent, instrument, and critical parameters, should be discussed and understood through technical teams between the originating lab and receiving lab.
2. The protocol must be defined and approved regarding parameters, acceptance criteria, time frame, and samples.
3. Deviation and investigation may be needed.
4. Closure and final report approval should take place.

The method transfer takes the form of a partial revalidation by the receiving lab and may include the following parameters, depending on the analytical procedure: system suitability, linearity, accuracy by recovery study, repeatability, QL/DL for related substance, and residual solvent tests.

For an ILQ, parameters may include system suitability, reproducibility, QL/DL for related substances, and residual solvent tests if QL/DL is not part of system suitability.

21.4.2 Potency

Typically, acceptance criteria for the potency assays in solid oral dosage forms are the same as that for intermediate precision in method validation. If there is any failure to meet the acceptance criteria, investigation will be conducted. For example, when the two results from two separately taken samples from the same batch are reliable but not in agreement with each other, drug product variation may be a cause. Product variation is not unusual for products like active coated tablets. In this case, it is necessary to prepare a composite sample in one lab and have this sample analyzed by both labs.

For the “drop” method (ie, preparing a sample by dropping certain units of tablets without grinding) for potency assays or separate analytical procedures for

CU testing, representative samples should be carefully evaluated in order to limit variation of the products. Another attention should be paid to purity of the reference standard if pretreatment of a reference standard such as the drying process is required. The predefined drying and testing procedure between two labs should be addressed.

21.4.3 Related substance assay

Typically, acceptance criteria should be followed based on an approved method validation report.

During ILQ practices, aged samples like the stability samples may be selected so that the samples contain impurities to aid in the evaluation of the data. If such samples are not available, the samples spiked with known impurities and stressed drug substances may be used where they are analyzed for recovery by both labs. In some cases, in order to ensure that each lab uses the same representative samples, a single laboratory is required to prepare the samples and the samples are then transferred to the second laboratory in a timely fashion.

21.4.4 Residual solvent assay

The acceptance criteria are similar to those defined in related substance assays except that samples spiked with solvents at the specification level should be used. It is most likely that the samples spiked with the solvent of interest are prepared by each individual lab and analyzed for recovery study. The absolute difference as generated by both labs in percent recovery should be no more than 20.0%, with 20.0% RSD for six replicate preparations from each lab.

21.4.5 Dissolution or release assay

Acceptance criteria should be established based on formulation (eg, immediate release or controlled/extended release) and sampling points (single or multiple). See chapter "Dissolution Testing of Solid Products" for more information on how to establish the acceptance criteria of dissolution tests for solid oral drug products. For immediate release products, there are at least two early sampling points for which percent release is less than 85% and one final sampling point. A total of 12 units should be tested in each individual lab. There are two acceptable approaches, direct comparison and model independent approach using similarity factor f_2 and difference factor f_1.

In general, the percent mean difference between both labs in percent release should be less than 10% for time points with less than 85% release, and no more than 5% for the final sampling point or Q-point.

The model independent approach using f_2 and f_1 can be used only if the criteria for f_1 and f_2 are met. The receiving lab is qualified if the f_1 value is within the range 0−15 and the f_2 value is within the range 50−100.

As per the FDA guidance for industry,[22] f_1 and f_2 are defined as follows:

$$f_1 = \left\{ \left[\sum_{t-1}^{n} |R_t - T_t| \right] / \left[\sum_{t-1}^{n} R_t \right] \right\} \times 100$$

where n is the number of time points, R_t is the dissolution value (in percent) of the reference (originating lab) at time t, and T_t is the dissolution wpvalue of the test (receiving lab) at time t.

$$f_2 = 50 \times \log \left\{ \left[1 + (1/n) \sum_{t-1}^{n} ((R_t - T_t)^2 \right]^{-0.5} \times 100 \right\}$$

The criteria to determine difference and similarity are as follows:

- 12 dosage units in each lab should be tested.
- Mean dissolution values at each time interval from both labs are used.
- Only one measurement should be used after 85% release.
- For mean values used, % RSD at the early time point should be less than 20%, and 10% for other time points.

21.5 CASE STUDIES

21.5.1 Case 1

21.5.1.1 Problem

A development Phase 1 batch of an API was analyzed for related substances. The HPLC test method was conducted using external standardization and using the API as the standard for any unknown impurities. For this purpose, the standard solution included a concentration of API equivalent to 0.10% of the API in the sample solution. The related substance specification for any unknown impurity was "Not more than 0.10%." The method had been validated to Phase 1 based on specified, identified impurities, and the QL was found to be 0.04%.

When one batch was analyzed in one laboratory (Lab 1), it was found to contain an unknown impurity with an estimated level of about 0.08%, which met the specification requirements. The same sample was then reanalyzed in another laboratory (Lab 2), and the reported level of the same impurity was 0.12%, which failed to meet the specification requirements.

21.5.1.2 Investigation

The out of specification (OOS) investigation conducted at Lab 2 confirmed that there were no analyst errors, and no assignable causes were identified. Retesting a new aliquot taken from the same sample confirmed the original result of 0.12%. The investigation was then extended to Lab 1, where again no analyst errors were identified, but retesting of a new aliquot taken from the same sample confirmed its original result of 0.08%.

The investigation was then extended to a closer examination of the chromatograms at both laboratories. Chromatograms were being monitored at a wavelength of around 230 nm, which corresponded with the maximum wavelengths for the API and each of the specified, identified impurities. As part of the further investigation, both laboratories repeated the analysis, but using DADs to check the spectra of the API and each impurity. It was verified that the maximum wavelengths for each known component was around 230 nm. For the unknown impurity, the maximum wavelength was around 223 nm, and the 230 nm reading actually coincided with a steep downslope of the spectrum. When both laboratories used 223 nm as the monitoring wavelength for this impurity, they then obtained agreeable results.

It should be noted that problems like this case can occur when chromatograms are being monitored using detector wavelengths that correspond to the edge of the spectrum for one or more components. The typical instrument specification of wavelength accuracy for UV detectors is ± 4 nm. The monitoring wavelength was set at 230 nm, but the true wavelength being monitored varies from 226 to 234 nm. Little difference for a component having its maximum absorbance at 230 nm is likely to be noted, since UV absorbance spectra of organic molecules normally show relatively broad peaks around 230 nm. However, this could lead to relatively large differences in the recorded absorbance values for any components that have peak edge at 230 nm.

To avoid this type of problem, it is recommended that the monitoring peak edge be eliminated wherever possible.

21.5.2 Case 2

21.5.2.1 Problem

During a validation study for a related substance test, the recovery was investigated by spiking with the impurities concerned. For one of those impurities, the recovery study was conducted using a solution at a concentration of 1 μg/mL. Although this level should be above the DL (which has not yet been verified in this validation study), no peak attributable to this impurity was observed in the chromatogram.

21.5.2.2 Investigation

The initial investigation verified that the spiking solution had been correctly prepared and all instrumental settings were correct, but no peak was generated from that solution. Further investigation could lead to a number of possibilities, discussed next.

21.5.2.2.1 Possibility 1

During method development, there may have been a change of monitoring wavelength to improve method sensitivity for another new component, but the analysts overlooked checking the effect of the change on other established components. It is very important during development to consider the impact of any changes on the overall analysis. In this circumstance, it may become necessary to have different monitoring wavelengths for different components to ensure adequate sensitivity for all components of interest.

It is usually recommended that the determinations of DL and QL be conducted as the first group of parameters during a validation study, which could have ensured the identification of this issue before further work was completed. In addition, these parameters are often required in order to design other portions of the validation program.

21.5.2.2.2 Possibility 2

The sample of the impurity or impurity standard that was used during development was from a different lot than that being used in the validation study, and the lots may not have the same quality. It is expected that API batches are synthesized under tightly controlled conditions in accordance with cGMP requirements. However, the same may not necessarily be true for impurities that are not the intermediates of the API and are being specially synthesized. So once an impurity of this type has been identified, it is very important to do the following:

1. Ensure that adequate controls are put into place so that we can be confident that we are really manufacturing the same compound each time. So it is important to put into place controls of the structural identity for different batches.
2. Ensure that the overall purities of different batches are kept reasonably constant.

21.5.3 Case 3

21.5.3.1 Problem

During the validation of precision and accuracy for the related substance test for a tablet product, a broad

peak was observed, which appeared in chromatograms after five sample injections. It accounts for a level of about 1.0%, which made quantitation of impurity peaks very difficult.

21.5.3.2 Investigation

The initial investigation showed no analytical errors, and the observed additional peak was genuine. This issue was unexpected, as the broad peak was not observed during method development and only noticed during the validation. The reason is likely due to sufficiently long runs applied during the validation, which would allow a component with strong retention to be eluted after five sample injections.

This additional component might be due to a tablet excipient, in which case it would not be a problem from the point of view of a new impurity, but the issue would need to be overcome to prevent interfering with this related substance test.

Alternatively, this peak might be due to a genuine unknown related substance, in which case, in accordance with ICH Guideline Q3B(R), it will need to be identified and then perhaps qualified, which could lead to delays in this product development program.

21.5.4 Case 4

21.5.4.1 Problem

While performing an accuracy experiment for the potency assay for a tablet, it was found that the recovery was 10% less than the theoretical value, although during method development, it was found that the sample extraction procedure had performed quite efficiently.

21.5.4.2 Investigation

During the investigation, it was discovered that a new batch of filters had been purchased for the validation study. Although a filter recovery study was completed satisfactorily with the previous batch, the new batch did not show a similar efficiency. Similar issues have also been experienced with HPLC autosampler vials when adsorption of an analyte on the surface occurs, which can happen in some cases.

21.5.5 Case 5

21.5.5.1 Problem

During a tablet stability study, it was observed that the assay values of the samples appeared to be declining for samples stored under accelerated conditions at 40°C/75% RH. There were no increases observed in degradation product peaks in the chromatograms.

21.5.5.2 Investigation

One possibility could be that the active ingredient degraded to products that were not being detected by the current analytical methodology. Normally, it would be expected that such a possibility would have been eliminated by conducting forced degradation (stress) studies prior to commencing the stability study, to ensure that the analytical methodology was suitable for monitoring the formation of tablet degradation products.

In this case, it was subsequently determined that under the accelerated storage conditions, the tablet matrix underwent some physical changes that reduced the efficiency of the extraction procedure during the sample preparation stage of the analysis. This type of issue is relatively common during development stages, and it is highly recommended that as part of an OOS investigation, the efficiency of the procedure for preparing analytical solutions be examined. Even though a method has been fully validated and the determination of accuracy using recovery studies has been deemed satisfactory, this can still be an issue. Validation studies are typically conducted using either freshly prepared product or product that has been stored at the suitable temperature conditions. They are not normally conducted using samples kept under accelerated storage conditions.

21.5.6 Case 6

21.5.6.1 Problem

During a stability study, trending analysis for the impurity profile demonstrates that the content of one known impurity at the 18M time point stored in 25°C/60% RH does not fall within the increasing trend over time.

21.5.6.2 Investigation

The product contains two active ingredients with more than 10 known impurities. The test method has been fully validated and implemented to monitor the impurity profile for release and stability testing for submission batches and market products over the years. It was observed that this specific known impurity increased over the storage time. Investigations indicated that the system suitability of the method met predetermined acceptance criteria. The reviews of all the chromatograms revealed that separation of this impurity from the main peak varies over the years. The problem is not resolved when a new column is used instead of using the aged column.

Upon further investigation, the problem could be solved if the performance of columns is screened and a good one selected from several new columns. Obviously,

this is impractical and not cost effective. Investigation then demonstrates that the resolution of these two peaks increases with the decrease of column temperature. However, the separation between other peaks was compromised at lower column temperatures. To solve the problem, the procedure with optimized column temperature was proposed to determine this impurity.

21.6 CONCLUSIONS

It is a regulatory requirement that all analytical methods that are used to generate data in a marketing authorization dossier be fully validated. It is generally considered acceptable in the industry to conduct phase-appropriate validation using the approaches discussed in this chapter as a product progresses through development. During the evolution of analytical methodologies, it is quite likely that methods will need to be applied in laboratories that did not participate in validation, and it is a cGMP requirement that the method be shown to function reliably under such circumstances via the applications of method transfers, as discussed in this chapter. Finally, regardless of how thoroughly validations and transfers are conducted, things will go wrong from time to time, and under such circumstances, it is very important to conduct thorough investigation and to be prepared to address a wide variety of root causes.

It is essential that clearly documented studies be conducted according to cGMP practices during analytical method development, validation, and method transfer.

References

1. International Conference on Harmonization Q2 (R1). Guideline: Validation of analytical procedures: text and methodology; 2005.
2. US Food and Drug Administration, Center for Drug Evaluation and Research. Guidance for industry: analytical procedures and methods validation for drugs and biologics; July 2015.
3. Ermer J, Miller JHMcB. *Method validation in pharmaceutical analysis: A guide to best practice*. Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA; 2005.
4. Swartz ME, Krull IS. *Analytical method development and validation*. New York, NY: Marcel Dekker; 1997.
5. US Food Drug Administration, Center for Drug Evaluation and Research. Guidance for industry, investigators, and reviewers: exploratory IND studies; January 2006.
6. US Food Drug Administration, Center for Drug Evaluation and Research. Guidance for industry: content and format of investigational new drug application (INDs) for phase 1 studies of drugs, including well-characterized, therapeutic, biotechnology-derived products; November 1995.
7. US Food Drug Administration, Center for Drug Evaluation and Research. Guidance for industry: INDs for phase 2 and phase 3 studies, chemistry, manufacturing, and controls information; May 2003.
8. European Medicines Agency, Inspections. CHMP/QWP/185401/2004, Guideline on the requirements to the chemical and pharmaceutical quality documentation concerning investigational medicinal products in clinical trials; December 16, 2004.
9. International Conference on Harmonization, 2002, The common technical document for registration of pharmaceuticals for human use: quality M4Q(R1).
10. USP 37, General Chapter <1225> Validation of compendial methods.
11. International Conference on Harmonization, ICH Q3A(R2). Impurities in new drug substances; 2006.
12. International Conference on Harmonization, ICH Q3B(R2). Impurities in new drug products; 2006.
13. International Conference on Harmonization Q3C(R3). Impurities: guideline for residual solvents; 2005.
14. Hauck WW, Foster T, Sheinin E, Cecil T, Brown W, Marques M, et al. Oral dosage form performance tests: new dissolution approaches. *Pharm Res* 2005;**22**(2):182–7.
15. Fortunate D. Dissolution method development for immediate release solid oral dosage forms "quick start guideline for early phase development compounds". *Dissolution Technol* 2005;**12**(3):12–14.
16. Lagace M, Gravelle M, Di Maso M, McClintock S. Developing a discriminating dissolution procedure for a dual active pharmaceutical product with unique solubility characteristics. *Dissolution Technol* 2004;**11**(1):13–17.
17. Qureshi SA. Developing discriminatory drug dissolution tests and profiles: some thoughts for consideration on the concept and its interpretation. *Dissolution Technol* 2006;**13**(4):18–23.
18. Brown CK, Chokshi HP, Nickerson B, Reed RA, Rohrs BR, Shah PA. Acceptable analytical practices for dissolution testing of poorly soluble compounds. *Pharm Technol* 2004;**12**:56–65.
19. Degenhardt OS, Waters B, Rebelo-Cameirao A, Meyer A, Brunner H, Toltl NP. Comparison of the effectiveness of various deaeration techniques. *Dissolution Technol* 2004;**11**(1):6–11.
20. Rohrs BR. Dissolution method development for poorly soluble compounds. *Dissolution Technol* 2001;**8**(3):6–12.
21. Viegas TX, Curatella RU, VanWinkle LL, Brinker G. Intrinsic drug dissolution testing using the stationary disk system. *Dissolution Technol* 2001;**8**(3):19–23.
22. USP30/NF25 <1092> The dissolution procedure: development and validation.
23. Food Drug Administration, Center for Drug Evaluation and Research. Guidance for industry: dissolution testing of immediate release solid oral dosage forms; 1997.
24. USP37 <711> Dissolution.
25. Gohel MC. Overview on chirality and applications of stereoselective dissolution testing in the formulation and development work. *Dissolution Technol* 2003;**10**(3):16–20.
26. Boehm G, Clark J, Dietrick, et al. The use of stratefied sampling of blend and dosage units to demonstrate adequacy of mix for powder blend. *PDA J Pharm Sci and Tech* 2003;**57**(2):64–74.
27. Khan MA, Kumar S, Jayachandran J, Vartak SV, Bhartiya A, Sinha S. Validation of a stability indicating LC method for aminodarone HCl and related substances. *Chromatographia* 2005;**61**(11/12):599–607.
28. Sajonz P, Wu Y, Natishan TK, McGachy NT, DeTora D. Challenges in the analytical method development and validation for an unstable active pharmaceutical ingredient. *J Chromatogr Sci* 2006;**44**:132–40.
29. Blanco M, Villar A. Development and validation of a method for the polymorphic analysis of pharmaceutical preparations using near infrared spectroscopy. *J Pharm Sci* 2003;**92**(4):823–30.

30. Sokoließ T, Köller G. Approach to method development and validation in capillary electrophoresis for enantiomeric purity testing of active basic pharmaceutical ingredients. *Electrophoresis* 2005;**26**:2330–41.
31. Korczynski MS. The integration of process analytical technologies, concurrent validation, and parametric release programs in aseptic processing. *PDA J Pharm Sci Technol* 2004;**58**(4):181–91.
32. Moffat AC, Trafford AD, Joe RD, Graham P. Meeting the international conference on harmonisation's guidelines on validation of analytical procedures: quantification as exemplified by a near infrared reflectance assay of paracetamol in intact tablets. *Analyst* 2000;**125**:1341–51.
33. Blanco M, Coello J, Iturriage H, Maspoch S, Pou N. Development and validation of a near infrared method for the analytical control of a pharmaceutical preparation in three steps of the manufacturing process. *Fresenius J Anal Chem* 2000;**368**:534–9.
34. Pappa H, Porter D, Russo K. Development of a new USP general information chapter: verification of compendial procedures. *Pharm Technol* 2006;**30**:164–9.
35. EP impurity 5.10.
36. USP 37 <1086> impurities in official articles.
37. International Conference on Harmonization, Q6A. New drug substances and new drug products: chemical substances (including Decision Trees); 1999.
38. International Conference on Harmonization Q1A(R2). Stability testing of new drug substances and products; 2003.
39. Zhou W, Yu K, Fang F, Carr G. Evaluation of HPLC-DAD for determination of co-eluting impurities in pharmaceutical analysis, submitted to publication.
40. Baertschi SW. *Pharmaceutical stress testing predicting drug degradation*. Boca Raton, FL: Taylor & Francis Group; 2005.
41. International Conference on Harmonization Q1B. Photostability testing of new drug substances and products; 1996.
42. US Food Drug Administration, Center for Drug Evaluation and Research. Reviewer guidance: validation of chromatographic methods; 1994.
43. USP37, <1010> Analytical data-interpretation and treatment.

This page intentionally left blank

CHAPTER

22

Statistical Design and Analysis of Long-Term Stability Studies for Drug Products

D. LeBlond

CMCStats, Wadsworth, IL, United States

This chapter is not intended to reiterate the very detailed and complete information that is already available in the research and textbooks on this subject. The literature in this area is vast, and an attempt to repeat it here would be a disservice to the reader. Instead, appropriate references are supplied, and the focus is on more recent statistical aspects, approaches, and recommendations that are important and useful and that are not discussed elsewhere. The material in Sections 22.1–22.3 is related to objectives, guidances, test methods, and data management. These sections can and should be profitably read and understood by formulation scientists, analytical chemists, and pharmaceutical development managers who have some understanding of the basic principles of experimental design. However, proper design and analysis of stability studies requires specific statistical expertise, and much of the material presented in the remainder of this chapter is technically quite complex and oriented to meet the needs and interests of professional statisticians.

22.1 STABILITY STUDY OBJECTIVES

The *stability* of a drug product is defined as the rate of change over time of key measures of quality on storage under specific conditions of temperature and humidity. A stability study should always be regarded as a scientific experiment designed to test certain hypotheses (such as equality of stability among lots) or estimate certain parameters (such as shelf life). Similar to any other scientific process, the outcome of a stability study should lead to knowledge that allows the pharmaceutical manufacturer to better understand and predict product behavior. Thus, a stability study is not merely a regulatory requirement, but also a key component in a process of scientific knowledge building that supports the continued quality, safety, and efficacy of a pharmaceutical product throughout its shelf life.

Understanding the stability of a pharmaceutical product (or any of its components) is important for proper quality design at many stages of the product life cycle. Table 22.1 lists some examples of pharmaceutical stability studies/analyses and their objectives.

Stability is intimately connected to many other key quality aspects of a drug product. For instance, interpretation of the rate of change of a key measure requires knowing the associated product release levels, recommended storage conditions, packaging, and stability acceptance limits. Proper interpretation of stability study data requires an understanding of the chemical-kinetic processes, the accuracy and precision of the associated test method, and the statistical limitations of the stability study experimental design. The statistical methodology used to achieve a given objective often depends on the quantity and quality of the data available and on the life-cycle stage of the product.

22.2 REGULATORY GUIDANCE

The regulatory aspects of drug product stability are governed by a number of interrelated guidance documents. In the United States, the European Union, and Japan, these guidances are provided through the International Conference on Harmonization of Technical Requirements for Registration of

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00022-4

TABLE 22.1 Applications of Stability Studies in Pharmaceutical Development

Product development stage	Objective
Chemical characterization	• Accelerated studies to define degradation pathways
Formulation development	• Establish retest period for active ingredient • Excipient/packaging selection and compatibility studies
Clinical studies	• Verify stability of clinical supplies.
Product registration	• Shelf life estimation • Release limit estimation • Determine process capability with respect to release or acceptance limits • Comparison of stability of clinical, development, and registration batches
Postapproval commitment	• Shelf life confirmation/extension with long-term studies • Annual stability monitoring
Marketed product	• Determination of predictive model from historical data • Shelf life extension • Assess risk of temperature excursions • Routine trending • Justification of scale-up, process, formulation, dosage strength, manufacturing site, packaging, or other postapproval changes • Establish equivalency with new formulations/packages • Annual stability reports

Pharmaceuticals for Human Use (ICH). Q1A-E[1–5] governs the stability testing of drug substances and products. Q2A-B[6,7] governs the validation of analytical methods used for (among other things) stability testing. Q3A-B[8,9] governs impurity levels in drug substances and products. Q6A-B[10,11] governs the acceptance criteria.

One must use caution when interpreting and using such regulatory guidances. Often, their individual scopes are somewhat narrow and may not always include the specific objectives of interest to the developer. There is no guarantee that following such guidances will lead to a manufacturable, approvable, high-quality product. Each individual guidance is established by a separate committee, and inconsistencies are inevitable. Blind adherence may have undesirable long-term consequences. Some common pitfalls are noted in the next discussion. A developer must always consider the specific study objectives and the approach taken should always be scientifically and statistically justified.

22.3 TEST METHODS AND DATA MANAGEMENT

Quantitative, precise, unbiased, stability-indicating test methods are essential components of stability experimentation. All information obtained from a stability study ultimately comes from the test methods used. No amount of statistical sophistication can compensate for a poor measurement system. The following recommendations are meant to promote the use of high-quality test methods and data management approaches:

- Use quantitative test methods that produce a result on a continuous measurement scale. These have higher information content than those that produce binary (yes/no), discrete (1, 2, 3, . . .), ordinal (low, medium, high), or categorical (A, B, C, . . .) responses.
- The validity of information obtained from a stability study ultimately depends on the test methods used. Where possible, the test methods used should be thoroughly characterized and validated prior to use. Partial validation may be acceptable for early development studies, but the uncertainties of the test method need to be accounted for in data evaluation. The validation should include the determination of bias and imprecision as a function of the true concentration level. The components of variance (replication, instrument, operator, day, lab, calibration run, etc.) should be identified and quantified. Such understanding is important in both study design (eg, assuring adequate sample size or study power) and analysis (eg, deciding whether to average replicated test values).
- Excessive rounding and binning of continuous measured values should be avoided, eg, the conversion of numerically measurable values to "below quantitation" because they are thought have unacceptable variability, or the rounding of results for minor components and degradation products to conform to specifications in regulatory guidelines (eg, ICH Q3A(R2) and Q3B(R2)). Such practices are common when reporting individual analytical results and may be required for the purpose of conforming to regulatory guidances when used in

this manner; however, when values such as stability testing results are used as input for further data analyses, overrounding and binning of test results lowers the data information content and distorts the measurement error structure. It should be the decision of the statistician performing the analysis of the data to round results appropriately after all data analysis has been completed. These abuses may limit the statistical procedures available for data analysis and lead to biased estimates or incorrect conclusions. All information contained in the original measurements should be used for efficient, sound decision making.

- Storage time should be well controlled. Regression procedures often assume that storage time is known exactly. Errors in the time scale may introduce bias and uncertainty that undermines the most accurate and precise response measurements. Stability studies typically follow predefined plans, in which sampling times are specified. However, for the purpose of data analysis, only the actual time at which a sample is taken is relevant; this must be accurately reported, especially if it deviates from the time specified in the plan. Company procedures should limit the time between sampling and analysis to as short a time as possible to prevent additional degradation; ideally, samples should be obtained immediately prior to analysis, even if this requires delaying sampling until after the time specified in the plan. Release test results (obtained at the end of manufacturing) should not be used in place of a true initial (zero time) test result when the release test date is substantially different from the starting date of a stability study. The storage conditions of the manufactured or sampled product should be appropriately controlled.
- Design stability databases with data analysis in mind. Statistical analyses may lead to approval of a longer shelf life, or they may provide the necessary information to better manage a product throughout its life cycle. Often statistical analyses are not performed or included in submissions because hand reentry of data is required, excessive reformatting must be performed, or there may be a delay in obtaining analytical results.
- Include a trained statistician on the study design and analysis team.[12]

Analytical data are obtained at great expense. The information contained within these data represents a proprietary advantage to the product developer. Thus, the analytical methods used to produce the data, as well as the computing/statistical methods used to extract information from them, should be of the highest possible quality.

22.4 MODELING INSTABILITY

22.4.1 Stability study variables

This section starts with shopping lists of key response, experimental, and controlled variables that often appear in drug product stability studies. These lists are by no means exhaustive. Then we will discuss how each item is typically incorporated into the kinetic model.

22.4.1.1 Responses

The response variable that is monitored over time should include, as appropriate, results from the physical, chemical, biological, microbiological, and key performance indicators of the dosage form. The potency level of the active ingredients and the levels of related degradants or impurities are always considered. In the case of instability due to degradation, mass balance should be accounted for. Only quantitative variables will generally be amenable to statistical analyses, as described later in this chapter.

Other critical or key quality attributes also need to be considered if they have acceptance limits. Dissolution may be especially critical for modified release products or products of poorly soluble compounds where bioavailability may change over time when dissolution is the rate-limiting step for drug absorption. Other responses that may be included in stability studies are color, moisture or solvent level, preservative, antioxidant, and pH.

22.4.1.2 Experimental fixed variables

Variables such as storage time, dosage strength, and packaging type are called *fixed* variables because the levels of such variables (eg, 3 months of storage or 300 mg) have specific meanings that are the same whenever that level is used in the study. Typically, except for storage time, fixed variables are categorical in nature and determined in advance by the experimenter.

Storage time is, of course, the primary experimental fixed variable always included in a stability study. ICH Q1A[1] recommends testing every 3, 6, and 12 months during the first, second, and subsequent years, respectively. Thus, for a typical 3-year study, testing would occur at 0, 3, 6, 9, 12, 18, 24, and 36 months. However, although storage times may be fixed by the study plan, they are actually measurable. Whenever and wherever possible, the exact measured elapsed time between the start of the study and the removal of a sample for analysis should be accurately reported, at least to the nearest whole day. Every effort should be made to minimize the elapsed time between when a sample is removed for analysis and the time at which the analysis is actually performed.

Other experimental fixed variables (also referred to as *factors* or *predictors*) that may be included as part of more complex studies include container type or size, closure type, fill size, desiccant type or amount, manufacturing site, batch size, and other covariates.

22.4.1.3 Experimental random variables

Variables such as test result replicate number or batch number are called *random* variables because the levels used (say replicate 1, 2, or 3) are not the same whenever they are applied; instead, these levels are assumed to represent experimental conditions drawn at random from a hypothetical infinite population of all possible replicates.

The ICH Q1A[1] guidance recommends that stability studies for new product registration include at least three batches of drug product (say batches 1, 2, and 3). If batch is considered a fixed variable, like dosage strength, then the estimated shelf life will reflect only the performance of the three batches in the study. If, on the other hand, batch is considered to be a random effect, like replicate, then the estimated shelf life will reflect the hypothetical infinite population of all future batches. To properly treat batch as a random variable requires the use of a so-called mixed regression model and specialized statistical software, such as the MIXED procedure in the SAS system.[13]

The shelf life estimate will generally be shorter if batch is regarded as random, but since the shelf life specification is meant to apply to all future batches, this may be more realistic. However, in most new product registrations governed by ICH Q1E,[5] where a minimum of three batches are available, batch is treated as a fixed variable and any inferences are strictly limited to the three batches on hand. This is done, quite simply, because data from only three batches are insufficient to project to the larger population of all future batches. This caveat must be kept in mind when interpreting the results of stability analyses.

22.4.1.4 Variable transformations

Sometimes the stability profile of a quality characteristic does not follow a straight line. Examples may include moisture level, which approaches equilibrium, or an intermediate degradation product, which follows a complex kinetic mechanism. In such cases, a nonlinear regression procedure should be considered. A description of nonlinear regression is beyond the scope of this chapter.

When the rate of change in the response measure is nearly constant over time and the change is monotonic, such that a transformation yields a nearly linear stability profile, linear regression (LR) may be used. LR provides a simple description that is easily grasped by nonstatisticians, requires few assumptions, and can be executed in widely available software. While the assumptions/approximations of LR are few, they must be satisfied for any estimates, inferences, or predictions made using LR to be valid. The assumptions of LR are:

1. The response is linearly related to the storage time.
2. There is no uncertainty in the storage time value.
3. Errors in the response measurements are normally distributed.
4. Errors in the response measurements are mutually independent.
5. The error variance in the response measurement is the same for all measurements.

Various statistical procedures that can be used to verify these assumptions are discussed in popular statistical texts.[14]

If the kinetic processes of instability are more complex, the profile may exhibit curvature. If a theoretical kinetic model is available, then nonlinear regression approaches (such as with the NLIN procedure of SAS) may be the best way to draw inferences, estimate shelf life, and make predictions from the stability data. In some cases, a nonlinear model can be linearized by transformation. For instance, the first-order kinetic model $y = A\exp(-kt)$ can be rewritten $\log(y) = \log(A) - kt$. Thus, a log transformation of the response, y, may improve conformance to assumption 1.

When the kinetic process is not well understood, various transformations of the response or time scales (or both) may be tried in an effort to allow LR to be used for analysis. It is important that the transformation be valid for all possible responses or time values. For instance, transformations such as $\log(y/(A-y))$ or $\log(t)$ are undefined for $y \geq A$ or $t = 0$, respectively.

Transformation of the time scale can have subtle effects on estimation efficiency but generally presents few statistical issues. However, transformation of the response scale may fundamentally change the error structure of the response measurements (assumptions 2–5). In favorable cases, a transformation may be found that will improve conformance to all five basic LR assumptions. However, if the transformation does not have a theoretical basis, extrapolations beyond the duration of the study must be made with extreme caution.

22.4.1.5 Controlled variables

Temperature and humidity are usually controlled during a stability study. The product label storage condition will dictate the conditions used in studies conducted to estimate shelf life. Table 22.2, taken from ICH Q1A,[1] gives typical conditions used.

For controlled room temperature products susceptible to moisture loss, lower relative humidity (RH) storage conditions may be appropriate (40% for long-term

TABLE 22.2 Temperature and Humidity Control of Stability Studies

Label storage	Long-term conditions	Intermediate conditions	Accelerated conditions
Controlled room temperature	25°C ± 2°C/60% ± 5% RH	30°C ± 2°C/65% ± 5% RH	40°C ± 2°C/75% ± 5% RH
Refrigerated	5°C ± 3°C		25°C ± 2°C/60% ± 5% RH
Frozen	−15°C ± 5°C		5°C ± 3°C/ambient RH
Minimum time period covered by data at submission	12 months	6 months	6 months

and intermediate and 15% for accelerated conditions). For controlled room temperature products stored in water impermeable containers, ambient RH might be appropriate for different temperature conditions.

While storage temperature and RH are quantitative variables that are often included in drug product stability studies, they are not generally included in a stability model. As specified by ICH Q1E,[5] each storage condition (ie, long-term, intermediate, or accelerated) is evaluated separately. An exception to this is with the analysis of accelerated stability studies, as described elsewhere in this book. Inclusion of storage condition variables as predictors in the stability model can be useful in judging the risk of temperature or humidity excursions, which can occur during the storage of finished drug products in warehouses or in a patient's environment. Such evaluations will not be discussed in chapter "Drug Stability and Degradation Studies."

22.4.2 A statistical model for instability

The design and analysis of a stability study requires the specification of a kinetic, predictive model for the instability of each key measure. A model should include mechanistic, experimental, and statistical aspects. That is, the model must take into account the physical or chemical mechanisms that result in changes in level over time. It must account for the fixed effects of variables whose levels are systematically varied as part of the stability study. Further, it must account for the statistical variation introduced by random variables whose levels are not specifically controlled, but vary in the study.

An understanding of the physicochemical mechanisms of instability of a drug product is an essential component of drug product life-cycle support. Knowing these mechanisms allows a developer to anticipate flaws, design an appropriate formulation and packaging system, and troubleshoot and support the product throughout its life cycle. The reader is directed to discussions elsewhere[15] for a discussion of kinetic mechanisms.

As indicated in ICH Q1A,[1] a linearizing transformation or use of an appropriate nonlinear model for the effect of storage time should be considered if the stability profile cannot be represented by a straight line. A thorough discussion of the important topic of physicochemical mechanisms is beyond the scope of this chapter. The chosen model for the effect of storage time should be scientifically and statistically justified.

Note to reader: The discussion thus far has been deliberately been kept at a level accessible to most nonstatisticians, but beyond this point, good familiarity with complex statistical methods is assumed. The advice of a professional statistician is strongly recommended before using any of the methods described next.

When the key measure changes only a small proportion (say, less than 15%) of its initial or potential level over its shelf life, a zero-order kinetic (straight-line) model is often found to be sufficient to describe the relationship between the level and storage time. Thus, a stability model often takes the form

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Zeta;}\gamma + \varepsilon \tag{22.1}$$

where

$\mathbf{y}$ is a column vector of test results.
β is a vector of fixed model parameters whose estimate is $\mathbf{b}$.
γ is a vector of random model parameters whose estimate is $\mathbf{g}$. It is typical to assume a multivariate normal distribution for γ.
ε is a vector of identically distributed and independent normal deviations with mean zero and true variance σ^2.

We take $\mathbf{X:Z}$ to be the design matrix for the stability study. $\mathbf{X}$ will contain a column of 1s for the intercept, a column of month values for storage interval (ie, the *actual* storage times, not the planned storage times), and columns of indicators for categorical fixed variables such as packaging or dosage strength. $\mathbf{Z}$ would contain columns of indicators for the levels of random variables such as batch. If batch is treated as a fixed variable, whose levels are specified in $\mathbf{X}$, and there are no other random variables, then the $\mathbf{Z}\gamma$ term is omitted.

22.5 LONG-TERM STABILITY STUDY DESIGN

22.5.1 Full and reduced designs

When a stability study includes all combinations of dosage strength, packaging, and storage times, the stability design is referred to as a *full-factorial* (or sometimes as just a *full* or *complete*) design. Conducting a full design on a new drug product with multiple strengths and packages can be prohibitively expensive and may stress a company's analytical resources beyond their limits. From a statistical point of view, when the number of study variables is large, a complete factorial may be unnecessary so long as certain assumptions can be made about the stability effects of the variables.

ICH Q1D[4] describes situations in which a reduced design can be applied without further justification, as well as some situations in which further justification will need to be provided. If a design deviates markedly from the principles of ICH Q1D[4], the protocol must be approved by the Food and Drug Administration (FDA) or other appropriate regulatory authority prior to the initiation of stability studies. Some additional clarification is provided in the literature.[16] In a reduced design, only a specific fraction of the possible combinations of dosage strength, packaging, and storage times are actually tested. ICH Q1D[4] refers to two general approaches to reduced designs: bracketing and matrixing.

22.5.2 Bracketing

In a bracketing approach, the sponsor relies on theory or past experience to identify a small number of variable combinations (say, of strength and packaging) that can be assumed to give worst-case stability. Often, these combinations will be extreme (eg, of active content, head space, moisture vapor transmission rate) and the sponsor is willing to estimate the product shelf life from a study of these worst cases alone. Bracketing assumes that the untested variable combinations will have equal or superior stability and therefore need not be tested at all. Bracketing requires a thorough understanding of the mechanisms of instability from theory or from studies on earlier development or clinical batches. Because bracketing makes strong assumptions about the underlying mechanism of instability, it is not applicable when dosage form formulations or package characteristics differ markedly. In a so-called pure bracketing design, the chosen combinations are tested at all time points. Often, those combinations not intended to be tested as part of the bracketing design are not even placed on stability.

Bracketing may be applied with no further justification across strength when different strengths have identical or closely related formulations (eg, capsules of different strengths made with different plug sizes from the same powder blend, or tablets of different strengths made by compressing varying amounts of the same granulation—formulations that differ only in minor excipients). Further justification should be considered when amounts of drug substance and excipients change in a formulation. When different excipients are used among strengths, bracketing is not generally applied.

Bracketing may be applied with no further justification across packages using the same container closure system where either container size or fill varies while the other remains constant. Further justification should be considered when the closure systems vary for the same container. Justification could include a discussion of relative permeation rates of the bracketed container closure system.

22.5.3 Matrixing

In a matrixing approach, the sponsor takes advantage of traditional principles of experimental design[17] to reduce study size without sacrificing statistical power, model structure, or parameter estimatability. Matrixing depends on choosing a balanced subset of the full factorial set of combinations that supports a predictive model that includes all main effects and critical interactions. Often, all combinations (even those not intended to be tested in the matrix design) are placed on stability should there be a need to revert to full testing.

Matrixing with respect to strength may be applied across strengths without further justification when the strength formulations are identical or closely related. Additional justification should be considered when different strengths differ in the drug substance content, excipients change, or different excipients are used. Matrixing across batches can be applied on batches that are made using the same process and equipment, or batches of drug substance. Matrixing across package size and fill is permitted when all packages use the same container closure system. Further justification should be provided when packages use different container closure systems, contact materials, suppliers, or orientations. Justification should be based on supportive data (eg, moisture vapor transmission rates or light protection for different containers).

Matrix designs can be complete (ie, all combination of factors are tested) or incomplete (ie, some combinations are not tested at all). In a complex design, combinations of strength and container size are tested and individual product batches are not tested in all strength and container size combinations. If the design

is broken during the course of the study, testing should revert to full testing through the proposed retest period or shelf life. Where testing exhibits moderate variability and moderately good stability, a matrix should be statistically justified.

Matrixing is not without risks. Highly fractional designs, involving factors other than time, generally have less precision in shelf life estimation and may yield shorter shelf life than a full design. With large variability and poor product stability, a matrix should not be applied. Techniques for comparing and assessing the statistical power of stability designs are discussed next.

22.5.4 Stability design generation

Before a reduced design is considered, assumptions should be assessed and justified. Then, as a starting point, either the design (including all possible combinations) or an appropriately bracketed subset is taken as the full design. Then reduced designs are obtained by matrixing the full design. The reduced designs may be compared with respect to the following criteria:

- Probability of justifying the desired shelf life as a function of study duration.
- Power to detect effects of experimental variables on stability. Any reduced design should retain the ability to adequately detect differences in stability.
- Balance (ie, each combination of factor levels is tested to the same extent) to ensure the orthogonality of model parameter estimates.
- The total number of tests required (or total study cost).
- Ergonomic spread of testing throughout the study to optimize analytical resources.

ICH Q1D[4] mandates full testing of all studied factor combinations at the beginning and end of the study, as well as at the time of submission. It also recommends that at least three time points for each studied factor combination be available at the time of submission (nominally at 12 months' storage). ICH Q1D[4] provides examples of study designs. These are not the only designs to be considered, but they illustrate many of the principles of balance, as well as the practical constraints. These are discussed next.

22.5.4.1 Matrixing on time points only

A key principle of fractional factorial design is to maximize testing of the extremes of continuous variables. In the case of storage time, this means that full testing is required at initial study completion and submission (typically 12 months). Thus reduced testing for a 36-month study can only be considered at five time points: 3, 6, 9, 18, and 24 months. Consider a study with six combinations: two strengths, with three batches per strength. Assume that a one-third reduction in testing is desired. Which 20 of the $5 \times 6 = 30$ test points should be tested? The principle of balance requires that:

- Each of the 2 strengths be tested 10 times ($2 \times 10 = 20$)
- Each of the 3 batches be tested Z times ($Z \times 3 = 20$)
- Each of the 5 time points be tested 4 times ($5 \times 4 = 20$)
- Each of the 6 strength × batch combinations be tested Y times ($Y \times 6 = 20$)
- Each of the 10 strength × time combinations be tested 2 times ($2 \times 10 = 20$)
- Each of the 15 batch × time combinations be tested W times ($W \times 15 = 20$).

Note that W, Y, and Z cannot be whole numbers. In fact, unless the number of tests is evenly divisible by 2 (strength), 3 (batch), and 5 (time), some loss of balance is inevitable. In this case, the lowest number of tests that allows balance is $2 \times 3 \times 5 = 30$, which does not allow for any testing reduction at all. Similarly, if a one-half reduction in testing was desired, which 15 of the 30 test points should be tested? Balance would require that the number 15 be evenly divisible by 2 (strength), 3 (batch), and 5 (time). Since 15 is not evenly divisible by 2, full balance is not possible in this case either. In the ICH Q1D[4] examples, the compromise made is to allow more testing on some batch, strength × batch, and batch × time combinations than others. However, the continuous time variable is robust to loss of balance because of the assumption of a linear change over time. Thus, the inevitable nonorthogonality in the Q1D[4] design examples is probably negligible.

An example illustrating complete and partial balance is as follows. Assume that the sponsor desired an analysis at 24 months (perhaps to justify shelf life extension). Then full testing would be required at 24 months, and matrixing would be on the four time points 3, 6, 9, and 18 months only. Thus, the full design would require (2 strengths) × (3 batches) × (4 times) = 24 tests. Since 12 is divisible by 2, 3, and 4, a completely balanced half-reduction is possible. A half-reduction would require 18 tests; however, 18 is only evenly divisible by 2 and 3, not by 4. Therefore, only a partially balanced design is possible if a half-reduction is desired. To identify which of the 12 or 18 tests to include, the mod arithmetic method of Nordbrock[18] can be used. The following steps illustrate this approach:

1. Assign code of S = 0 or 1 for each strength.
2. Assign a code of B = 0, 1 or 2 for each of the three batches within each strength.

3. Assign a code of T = 0, 1, 2, or 3 for the time points 3, 6, 9, or 18 months, respectively.
4. For each of the 24 possible strength × batch × time combinations,
   a. For a half-fold reduction, test only combinations where S + B + T mod 2 = 0.
   b. For a one-quarter-reduction, test only combinations where S + B + T mod 3 = 0 or 1.

Tables 23.3–23.5 show the testing schedule at these time points for these two cases.

The particular ½ or ⅓ fraction selected in Tables 22.3 or 22.4 are, of course, only one of two or three possible fractions. For instance, one could have decided to test S + B + T mod 2 = 1 in Table 22.3 or S + B + T mod 3 = 0 or 2 in Table 22.4 instead. The sum S + B + T can be generalized to $n \times S + B + m \times T$, where n and m are constants other than 1, in an attempt to find a design that achieves the desired balance. The same principles can be extended to more complex situations. The example in ICH Q1D[4] involves a product with 3 strengths, 3 packages, 3 batches, and 5 time points. It is typical to first identify a set of balanced (or approximately balanced) time vectors. Then the various strength × package × batch combinations

TABLE 22.3 Illustration of Use of Indicator Variable Summation to Find a Matrix Design on Time Points for a Product With Two Strengths

		S + B + T			
Strength	**Batch**	**3 Months (T = 0)**	**6 Months (T = 1)**	**9 Months (T = 2)**	**18 Months (T = 3)**
Low (S = 0)	1 (B = 0)	0	1	2	3
	2 (B = 1)	1	2	3	4
	3 (B = 2)	2	3	4	5
High (S = 1)	1 (B = 0)	1	2	3	4
	2 (B = 1)	2	3	4	5
	3 (B = 2)	3	4	5	6

TABLE 22.4 Example of a Balanced Half-Reduction Matrix Design on Time Points for a Product With Two Strengths

		S + B + T mod 2			
Strength	**Batch**	**3 Months (T = 0)**	**6 Months (T = 1)**	**9 Months (T = 2)**	**18 Months (T = 3)**
Low (S = 0)	1 (B = 0)	0	1	0	1
	2 (B = 1)	1	0	1	0
	3 (B = 2)	0	1	0	1
High (S = 1)	1 (B = 0)	1	0	1	0
	2 (B = 1)	0	1	0	1
	3 (B = 2)	1	0	1	0

0—test; 1—do not test.

TABLE 22.5 Example of a Partially Balanced One-Third-Reduction Matrix Design on Time Points for a Product With Two Strengths

		S + B + T mod 3			
Strength	**Batch**	**3 Months (T = 0)**	**6 Months (T = 1)**	**9 Months (T = 2)**	**18 Months (T = 3)**
Low (S = 0)	1 (B = 0)	0	1	2	0
	2 (B = 1)	1	2	0	1
	3 (B = 2)	2	0	1	2
High (S = 1)	1 (B = 0)	1	2	0	1
	2 (B = 1)	2	0	1	2
	3 (B = 2)	0	1	2	0

0 or 1—test; 2—do not test.

are then assigned in a balanced (or approximately balanced) way to each of the vectors. The example in ICH Q1D[4] establishes the following testing vectors:

T1: {0, 6,9,12,18,24,36}
T2: {0,3, 9,12, 24,36}
T3: {0,3,6, 12,18, 36}

T1 calls for testing at 7 total time points, but T2 and T3 only 6. However, balance is achieved across the time points 3, 6, and 9. Because of the assumption of linear trend across time, lack of balance does not lead to serious degradation of estimation efficiency. A desirable feature of T1, T2, and T3 as given here is that each of the time points 3, 6, 9, 18, and 24 is represented twice. Thus, if the vectors are evenly spread across the combinations, a one-third-reduction of testing at these time points would be realized. The assignments of the vectors to each of the $3 \times 3 \times 3 = 27$ combinations is illustrated in Table 22.6.

22.5.4.2 *Matrixing on both time points and other variables*

Continuing the ICH Q1D matrix example, one can matrix further against batch, strength, and package by omitting the testing based on the value of B + S + P mod 3. In the ICH Q1D Table 3b example, the criterion for elimination depends on batch, as described in Table 22.7.

The resulting approximately balanced matrix provides about a 5/9 reduction in the amount of testing at the 3-, 6-, 9-, 12-, 18-, and 24-month time points. These examples are provided merely for illustration. Other matrixing designs are possible, and some may well be superior to those given. Once candidate designs are generated by the above-mentioned methods—or by other standard experimental design techniques—it is best to perform an assessment of their balance and operation feasibility. Those, which pass these assessments, may be examined further by the statistical methods described in the next section.

The level of reduction that would be acceptable to a regulatory agency such as the FDA can be addressed by examining the power of the design to detect differences in stability among the different batches, strengths, or packages and the precision with which shelf life can be estimated. These characteristics are compared between the candidate matrix design and the full design. If a candidate matrix design sacrifices little relative to the full design, then it should be acceptable from a regulatory point of view.

22.5.5 Comparing stability designs

Nordbrock[19] provided a framework for comparing stability designs with respect to their power to detect slope (stability) differences. This same framework can be applied to a comparison of the probability (power) of meeting a desired shelf life claim among competing stability designs.

22.5.5.1 *Review of preliminary statistical concepts*

Let R and H be mutually independent random variables with distributions $R \sim N(\delta,1)$ and $H \sim \chi^2(dfe)$. Then, by definition,

$$\frac{R}{\sqrt{H/dfe}} = W \sim \text{nct}(dfe, \delta) \tag{22.2}$$

where nct(dfe,δ) represents a random variable distributed as noncentral t with degrees of freedom dfe and noncentrality parameter δ. Further, let pnct($Q \mid dfe$,δ) be the cumulative distribution of the noncentral t random variable W relative to fixed Q such that

$$\text{Prob}\{W \le Q\} = \text{pnct}(Q|dfe, \delta) \text{ and } \text{Prob}\{W > Q\} = 1 - \text{pnct}(Q|dfe, \delta) \tag{22.3}$$

Noncentral t distribution functions are present in software packages such as SAS or R. The following approximations may be useful in implementing these calculations in packages that have a cumulative noncentral F (pf()), but not a noncentral t:

TABLE 22.6 Using Mode Arithmetic to Matrix on Time Points Only for the Example of Table 3a in ICH Q1D

	B + S + P mod 3								
Strength:	**Low (S = 0)**			**Med. (S = 1)**			**High (S = 2)**		
Package:	A (P = 0)	B (P = 1)	C (P = 2)	A (P = 0)	B (P = 1)	C (P = 2)	A (P = 0)	B (P = 1)	C (P = 2)
Batch 1 (B = 0)	0	1	2	1	2	0	2	0	1
Batch 2 (B = 1)	1	2	0	2	0	1	0	1	2
Batch 3 (B = 2)	2	0	1	0	1	2	1	2	0

0—use T1; 1—use T2; 2—use T3.

TABLE 22.7 Using Mod Arithmetic to Include Matrixing Against Design Variables in the Example of Table 3b of ICH Q1D

Batch	Do not test if B + S + P mod 3 =
1	2
2	1
3	0

$$\mathrm{pt}(t, dfe, \delta) = \mathrm{pf}(t^2, 1, dfe, \delta^2) + \mathrm{pt}(-t, dfe, \delta) \sim$$
$$= \mathrm{pf}(t^2, 1, dfe, \delta^2) \text{ for } \delta >> 0$$

In some software packages, algorithms for noncentral t do not converge when δ, dfe, or both are large. In those cases, the following approximation may be used:

$$\mathrm{pnorm}(t - \delta) \cong \mathrm{pt}(t, dfe, \delta)$$

with pnorm(z) the standard normal distribution function.

Now consider a LR to the fixed model:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{22.4}$$

with fixed model parameters $\boldsymbol{\beta}$ and $\boldsymbol{\varepsilon} \sim N(0,\sigma^2)$. In applying the concepts here, it is critical to understand the structure of $\boldsymbol{\beta}$ and $\mathbf{X}$. As an example, consider a stability design that includes measurements (y vector) at a series of time (T) points on three batches (B), each of three strengths (S), with three possible package (P) options. Assume that the design will support estimation of all main effects and certain interactions (ie, others known to be unimportant). In a high-level program like SAS, the model might be expressed as

$$\mathrm{Y} = \mathrm{intercept\ T\ B\ S\ P\ T*B\ T*S\ T*B*S}$$

The column vector $\boldsymbol{\beta}$ will be composed of one subvector for each model term. The length of each subvector is 1 for the continuous variables intercept and T. For categorical main effects B, S, and P, the subvector length will be number of categorical levels for that variable minus 1; therefore, the length will be 2 for each main effect. For interactions, the vector length will be the product of the vector lengths of each of the component main effects in the interaction. Working across in this model order, we see the length of β will be 1 + (1 + 2 + 2 + 2) + (2 + 2) + (4) = 16. Thus

$$\boldsymbol{\beta}' = (\beta_{\mathrm{int}}, \beta_T, \beta_{B1}, \beta_{B2}, \beta_{S1}, \beta_{S2}, \beta_{P1}, \beta_{P2}, \beta_{TB1}, \beta_{TB2}, \beta_{TS1}, \beta_{TS2}, \beta_{TB1S1}, \beta_{TB1S2}, \beta_{TB2S1}, \beta_{TB2S2})$$

Each element of $\boldsymbol{\beta}$ is an effect attributable to the presence of certain levels or combinations of levels of the variables T, B, S, and P.

The $\mathbf{X}$ matrix in this example will have 16 columns corresponding to each of the coefficients in β, and one row for each result in y. The intercept column in $\mathbf{X}$ will have a 1 in each row, and the T column in $\mathbf{X}$ will have the numerical storage time t at which the corresponding test result in y was obtained. For categorical main effects, the $\mathbf{X}$ columns corresponding to B, S, and P will contain codes that specify the level of each corresponding to the respective y test result. A convenient 0/1 coding is as follows (Table 22.8): in which the presence of a 0 in the first or second element indicates that the corresponding factor level is not acting to produce the corresponding result. If both the first and second elements are 0, then the third level must be active.

Subvector codes for interactions consist of the Kronecker product of the subvectors of the component main effects. Thus, the 2 T*B columns in $\mathbf{X}$ corresponding to a y result obtained at 3 months on Batch 2 would be encoded as $3 \otimes (0\ 1) = (0\ 3)$. The four T*B*S columns in $\mathbf{X}$ corresponded to a y result at 6 months on Batch 1, middle strength would be $6 \otimes (1\ 0) \otimes (0\ 1) = (0\ 6\ 0\ 0)$. In this way, the entire $\mathbf{X}$ matrix is constructed. Consider the y result corresponding to the 9-month test on batch 2 of low strength packaged in container type B. The corresponding row of the $\mathbf{X}$ matrix is, then,

$$X = \begin{pmatrix} \cdots \\ 1, 9, 0, 1, 1, 0, 0, 1, 0, 9, 9, 0, 0, 0, 9, 0 \\ \cdots \end{pmatrix}.$$

The estimates of interest in a stability analysis generally consist of linear functions of the elements of $\boldsymbol{\beta}$. Functions of interest might be slopes or differences between slopes for specific combinations of design variables, or averages across certain design variables. A linear function of $\boldsymbol{\beta}$, say $S = \mathbf{c}'\boldsymbol{\beta}$, will have an estimate $\hat{S} \sim N(S, \sigma_S^2)$ with a sampling variance as follows:

$$\sigma_S^2 = \mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}\sigma^2, \tag{22.5}$$

TABLE 22.8 Coding of Design Variable Levels in the Construction of the $\mathbf{X}$ Matrix

Subvector code	B Interpretation	S Interpretation	P Interpretation
(1 0)	Batch 1	Low strength	Package A
(0 1)	Batch 2	Middle strength	Package B
(0 0)	Batch 3	High strength	Package C

based on degrees of freedom *dfe* where

$$dfe = \text{number of rows in } \mathbf{X} - \text{number of columns in } \mathbf{X} \tag{22.6}$$

The dimensions of the contrast column vector **c** will be identical to that of β. The elements of **c** will depend on which of the corresponding elements of β are active in defining a specific linear function of interest.

Finally, note that if $\hat{\sigma}_S^2$ is a sample estimate of σ_S^2 based on *dfe* degrees of freedom, then its sampling distribution is given by

$$\frac{\hat{\sigma}_S^2}{\sigma_S^2} \cdot dfe \sim \chi^2(dfe). \tag{22.7}$$

22.5.5.2 *Power to detect slope differences*

Let S represent a slope difference that is of interest, and consider a statistical test of H0: $S = 0$ against an alternative Ha: $S > 0$. H0 will be rejected at the 0.05 level of significance if

$$\frac{\hat{S}}{\hat{\sigma}_S^2} = \frac{\hat{S}/\sigma_S^2}{\hat{\sigma}_S^2/\sigma_S^2} > qt(0.95, dfe) \tag{22.8}$$

Now, under the assumption that $S = \Delta$,

$$\frac{\hat{S}}{\sigma_S^2} \sim N\left(\frac{\Delta}{\sigma_S^2}, 1\right) \tag{22.9}$$

Combining Eqs. (22.2), (22.3), (22.6)–(22.8), we may state that

$$\begin{aligned}\Pr\{\text{H0 rejected}|\Delta\} &= \Pr\left\{\frac{\hat{S}}{\hat{\sigma}_S^2} > qt(0.95, dfe)|\Delta\right\} \\ &= 1 - \text{pnct}\left(qt(0.95, dfe), dfe, \frac{\Delta}{\sigma_S}\right)\end{aligned} \tag{22.10}$$

Eq. (22.10) may be evaluated as a function of Δ to provide the operating characteristics of the statistical test. Since, by Eqs. (22.5) and (22.6), *dfe* and σ_S^2 depend only on **X**, which in turn is determined by the stability design, the operating characteristics can be compared for various designs and storage times as an aid in study planning.

As an example, consider three matrix designs presented in ICH Q1D[4] and the model as described previously. Let *S* be a slope difference between medium and low (medium minus low) strength averaged across all lots. Then the corresponding contrast vector is

$$\mathbf{c}' = \left(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, -1, +1, -\frac{1}{3}, +\frac{1}{3}, -\frac{1}{3}, +\frac{1}{3}\right)$$

so that

$$\mathbf{c}'\boldsymbol{\beta} = \beta_{TS2} - \beta_{TS1} + \frac{\beta_{TB1S2} - \beta_{TB1S1} + \beta_{TB2S2} - \beta_{TB2S1}}{3}$$

Assume that the stability potency analysis for submission will occur at the 12-month time point. Let $\sigma = 1\%$LC, and let the true slope difference, Δ, vary between 0 and 0.25%LC/month. Fig. 22.1 shows the power (as given by Eq. (22.10)) for various true slope differences. Note that very little power is lost in going from the full design (135 tests required) to the matrix design on time only (ICH Q1D[4] Table 3a; 108 tests required). However, considerable power is lost when matrixing includes time, strength, batch, and packaging (ICH Q1D[4] Table 3b; 72 tests required).

22.5.5.3 *Probability of achieving a shelf-life claim*

Consider a stability-determining test, such as potency, with a lower acceptance limit, *L*. Let *D* be the desired shelf for the product. Further assume that the product will have a true release potency (intercept) of *I*, a true slope of *S* and a true analytical standard deviation of σ_S.

The regression slope estimate of *S* (ie, $\hat{S}$) will have a sampling distribution that can be defined by

$$\frac{\hat{S} - S}{\sigma_S} + \delta \sim N(\delta, 1) \tag{22.11}$$

Combining Eq. (22.11) with Eq. (22.7) and comparing with Eqs. (22.2) and (22.3), we see that

$$\Pr\left\{\frac{\frac{\hat{S}-S}{\sigma_S} + \delta}{\hat{\sigma}_S/\sigma_S} > Q\right\} = 1 - \text{pnct}(Q, dfe, \delta). \tag{22.12}$$

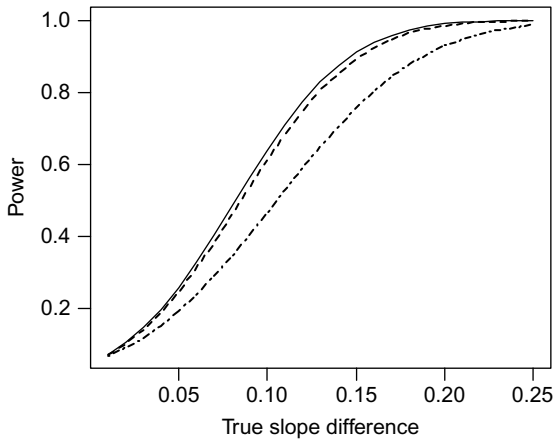


FIGURE 22.1 Comparison of power to detect slope differences at the 12-month time point for three stability designs in ICH Q1D. Solid line—full design; dashed line—matrix on time points only; dot-dashed line—matrix on both time points and design variables.

Now for a test with a lower limit, the desired shelf-life claim will be achieved whenever

$$I + D\cdot(\hat{S} - \hat{\sigma}_S \cdot qt(0.95, dfe)) > L$$

or

$$\frac{\hat{S} + (I-L)/D}{\hat{\sigma}_S} > qt(0.95, dfe)$$

or

$$\frac{\frac{\hat{S}-S}{\sigma_S} + \frac{(I-L)/D + S}{\sigma_S}}{\hat{\sigma}_S/\sigma_S} > qt(0.95, dfe)$$

Comparing this with Eq. (22.12), we see that

$$\Pr\{\text{Achieving desired shelflife claim}\} = 1 - \text{pnct}\left(qt(0.95, dfe), dfe, \frac{(I-L)/D + S}{\sigma_S}\right) \quad (22.13)$$

Note that a similar derivation in the case of an upper acceptance limit would yield

$$\Pr\{\text{Achieving desired shelflife claim}\} = \text{pnct}\left(qt(0.05, dfe), dfe, \frac{(I-L)/D + S}{\sigma_S}\right) \quad (22.14)$$

Eqs. (22.13) or (22.14) may be evaluated as a function of I, S, and σ_S to provide the operating characteristics of the shelf-life-estimation goal. Hypothetical I, S, and σ_S may be taken from preliminary stability and analytical studies. Since by Eqs. (22.5) and (22.6), dfe and σ_S^2 depend only on $\mathbf{X}$, which in turn is determined by the stability design, the operating characteristics can be compared for various designs and storage times as an aid in study planning.

As an example, consider three matrix designs presented in ICH Q1D and the model as described previously. Let S be defined as the slope that represents the stability for batch 2 of the low-strength formulation (and any package). The corresponding contrast vector is, then,

$$\mathbf{c}' = (0,1,0,0,0,0,0,0,0,1,1,0,0,0,1,0)$$

so that

$$\mathbf{c}'\boldsymbol{\beta} = \beta_T + \beta_{TB2} + \beta_{TS1} + \beta_{TB2S1}$$

Assume that the stability potency analysis for submission will occur at the 12-month time point. Let $\sigma = 1\%$LC, the lower acceptance limit for potency, L, be 90%LC, the true initial potency level, and I, be 100%LC, and let the true slope, S, for the combination of interest vary between -0.4 and -0.1%LC/month. Fig. 22.2 shows the probability of meeting a shelf life claim of $D = 24$ months (as given by Eq. (22.13)) for various true slopes. Note that very little risk is seen in going from the full design (135 tests required) to the matrix design on time only (ICH Q1D Table 3a; 108 tests required). However, considerable risk is encountered when matrixing includes time, strength, batch, and packaging (ICH Q1D Table 3b; 72 tests required) (Fig. 22.2).

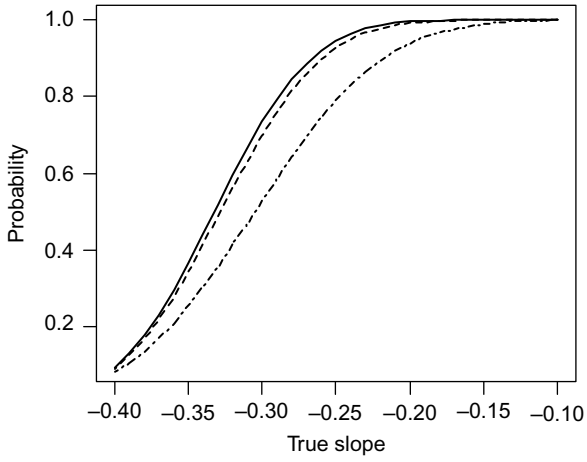


FIGURE 22.2 Comparison of probability of achieving a 24-month shelf-life claim at the 12-month time point for three stability designs in ICH Q1D. Solid line—full design; dashed line—matrix on time points only; dot-dashed line—matrix on both time points and design variables.

22.5.5.4 *Implementation in R*

The R language[20] provides a good matrix computation and graphics platform for making stability design comparisons of power and probability as described previously. The design matrix **X** (up to the time of proposed analysis) may be created manually in a Microsoft Excel spreadsheet and imported as a comma-separated text file to implement the power and probability calculations illustrated in the following R code:

```
XPXINV<-solve(t(X)%*%X) # information matrix
dfe<- dim(X)[1]-dim(X)[2]
var.e<- 1 # hypothesized analytical variance

# contrast vector for slope differences
c<-matrix(c
(0,0,0,0,0,0,0,0,0,0,-1,1,-0.334,0.333,

        -0.334,0.333),ncol=1)

# hypothesized true slope differences
DELTA<-seq(from=0.01,to=0.25, by=0.01)
var.c<-var.e*t(c)%*%XPXINV%*%c;var.c
# power of detection of a slope difference
power<-1-pt(qt(0.95,dfe),dfe,DELTA/sqrt(var.
c))
# contrast vector for individual combination slope
estimate
a<-matrix(c(0,1,0,0,0,0,0,0,0,1,1,0,0,0,1,0),
ncol=1)
# hypothesized true slopes
S<-seq(from=-0.1, to=-0.4, by=-0.01)
var.a<-var.e*t(a)%*%XPXINV%*%a;var.a
# prob of achieving shelf life goal
I<-100 # Initial potency level
L<-90 # Lower specification for dating
D<-24 # Desired dating (based on 12 month data)
prob<-1-pt(qt(0.95,dfe),dfe,(S+(I-L)/D)/sqrt
(var.a))
```

22.6 DETERMINATION OF SHELF LIFE

22.6.1 Definition of shelf life

A definition of *shelf life* for a given characteristic that is consistent with ICH Q1E[5] is the longest storage period for which there is at least 95% confidence that the lot mean level of the characteristic remains within its acceptance range. We must translate this definition into a procedure that we can apply for estimating the shelf life. The procedure used will depend on the type of data available and the statistical model employed. When the shelf life is estimated using LR, this definition is consistent with the storage period at which the one-sided upper or lower 95% confidence bound for the lot mean level intersects the product acceptance limit.

By this definition, the shelf life is clearly intended to control the mean level of critical characteristics of product lots. Because of analytical variance and content nonuniformity, even when the lot mean level of a characteristic is within a given specification, the measured levels of individual tablets may not be. The occurrence of individual out of specification (OOS) test results in lots tested annual stability monitoring can lead to a costly laboratory investigation. Thus, as part of a shelf-life justification, a sponsor may also want to evaluate the probability of individual failures.

Because this definition of shelf life does not control the spread of individual results, alternative definitions are proposed from time to time based on prediction or tolerance bounds.[21]

22.6.2 Model pruning

The first step in shelf-life estimation consists of pruning the full model to remove statistically nonsignificant terms. This process should follow the usual stepwise regression elimination that maintains model hierarchy. Generally, model terms that contain batch variables are eliminated at the $P > 0.25$ significance level, and other terms are eliminated at the more traditional $P > 0.05$ significance level. A good overview of the model building and fitting process is described by Chow and Liu[22] and by Shao and Chow.[23] The statistical aspects of the analysis for a fixed effects model are discussed by Fairweather et al.[24] and by Chen et al.[25]

When a mixed-effects model is used, the same pruning principles may be employed. Analysis of mixed effects stability models is described by Chen et al.[26] and by Chow and Shao.[27]

A set of SAS macros was made available from the FDA[28] that provided an analysis of the simple fixed-batch case with no other factors such as strength or packaging. However, this set of macros is not recommended because it is too limited in scope. It is not entirely compatible with the new ICH Q1E[5] guidance, nor with current versions of the SAS package. The MIXED procedure of SAS permits analysis of a mixed model (Eq. (22.1)) and is an excellent platform for shelf-life estimation. Therefore, examples of proc MIXED code are provided next.

22.6.2.1 *Simple ANCOVA for the fixed batch case*

In the simplest situation, stability data are available for at least three batches of a single strength and stored in a single package type. A good basic reference to ANCOVA is Brownlee.[29] The process begins with preliminary sequential *F*-tests for equality of degradation rates (slopes) and initial values (intercepts) among batches (Table 22.9).

Let us call the *P*-values of the Batch and Month*Batch tests *P*(B) and *P*(M*B), respectively. According to ICH Q1E, the final model is selected according to the following rules:

Select common intercept common slope (CICS) if $P(B) > 0.25$ and $P(M*B) > 0.25$
Select separate intercept common slope (SICS) if $P(B) < 0.25$ and $P(M*B) > 0.25$
Select separate intercept separate slope (SISS) otherwise.

An aspect of this model selection algorithm that is unique to the ICH Q1E[5] guideline is the use of a critical *P* value of 0.25 instead of the usual $P = 0.05$. The FDA has indicated that this should be done in order to protect against accidentally obtaining a favorable model (model 1 CICS generally leads to longer product dating) by the statistical accident of getting favorable data when in fact the real behavior is less favorable.[30] This is still a controversial aspect of the FDA/ICH procedure, but as this has been used for well over 15 years, it has become standard practice in industry.

The ANCOVA can be implemented simply using the MIXED procedure of SAS and will be illustrated using the data in Appendix A. Assume that the data for Batches 4, 5, and 8 are analyzed together (ie, *pooled*, in the drug stability vernacular[31]) and read into a SAS

TABLE 22.9 Simple ANCOVA Sequential *F*-Tests

Code name for the ANCOVA *F*-test	What is tested?
Batch	Differences among batch intercepts, assuming that batches have a common slope
Month*Batch	Differences among batch slopes, assuming that batches may have different intercepts

TABLE 22.10 Type 1 Tests of Fixed Effects

Effect	Num DF	Den DF	*F* Values	Pr > *F*
Month	1	18	100.99	<0.0001
Batch	2	18	72.12	<0.0001
Month*Batch	2	18	1.96	0.1704

data set with the variables Batch, Month, and Potency. The following SAS code:

```
proc mixed;
  class batch;
  model Potency = Month Batch Month*Batch/htype = 1;
run;
```

results in the output in Table 22.10.

The last column gives the P(B) and P(M*B) values on the second and third lines respectively. Since P(M*B) < 0.25, a SISS model is required to estimate the shelf life.

The reader may verify that, for the data in Appendix A, analysis of batches 3, 4, and 5 will lead to the SICS model and analysis of batches 2, 5, and 7 will lead to the CICS model.

22.6.2.2 *Model pruning for more complex studies*

Whether a full factorial or a reduced design is used, data should be treated in the same manner as data from a full design study. As an example, if the stability study includes multiple strength and packaging levels, and the design supports estimation of interactions to the second order, the following SAS code might be appropriate:

```
proc mixed;
  class batch strength package;
  model Potency = Batch Strength Package
                  Batch*Strength      Batch*Package
Strength*Package
                  Month  Month*Batch  Month*Strength
Month*Package
            /htype = 3;
run;
```

A model pruning procedure similar in spirit to the simple ANCOVA analysis given here should be used (eg, stepwise regression) to test for the statistical significance of model terms containing the various design variables. Model terms are eliminated in a stepwise manner based on an objective criterion, such as their corresponding Type III F test significance. Preliminary statistical tests for terms or interactions involving the Batch variable should be conducted at a level of significance for a rejection of 0.25. Statistical tests for main effects of matrixed or bracketed variables other than Batch should be conducted using the traditional level of significance for rejection of 0.05.

22.6.3 Simple fixed batch case

If the purpose of the stability analysis is to estimate stability parameters or predict performance of only those batches included in the study, the batch variable may be considered fixed. In this case, the **Z** matrix in Eq. (22.1) is null and the stability model is the usual fixed linear model:

$$y = \mathbf{X}\beta + \varepsilon. \tag{22.15}$$

where the **X** matrix encodes the levels of fixed effects, such as batch, strength, packaging, and storage time.

The estimated product shelf life depends on the final stability model. Assuming that model pruning has already been done, **X** will include columns only for effects and interactions deemed statistically significant whose estimation can be supported by the data. We consider here the single-strength, single-package case. Shelf-life estimation requires that the upper (or lower) one-sided 95% confidence bounds for the mean regression lines for each batch be calculated. The SAS MIXED syntax for the SISS model is given here:

```
proc mixed;
class batch;
  model Potency = Month Batch Month*Batch / Solution
alphap = 0.1
                                  outpm = outpm; run;
```

In SAS, confidence bounds are easily obtained for time points not studied by including missing values for desired months in the incoming data set. For the other models, the following SAS code should be substituted:

SICS: Potency = `Month Batch`

CICS: Potency = `Month`

The resulting outpm SAS data set will contain a listing of the data, as well as the appropriate predicted and confidence bounds for each time point in the incoming data set. Missing values may be included for time points of interest in order to estimate the shelf life. As an illustration, an analysis of Batches 4, 5, and 8 of the data in Appendix A will provide an `outpm` data set containing the following:

```
batch    Month    Lower
BATCH4    56    90.8585
BATCH4    57    90.6146
BATCH4    58    90.3708
BATCH4    59    90.1268 <--
BATCH4    60    89.8829
BATCH5    41    90.8015
BATCH5    42    90.5517
BATCH5    43    90.3017
BATCH5    44    90.0518 <--
```

```
BATCH5    45    89.8017
BATCH5    46    89.5516
BATCH5    47    89.3014
BATCH5    48    89.0512
BATCH8    25    90.9432
BATCH8    26    90.5079
BATCH8    27    90.0723 <--
BATCH8    28    89.6365
BATCH8    29    89.2004
```

The shelf life for a given batch is the maximum time (rounded to the nearest month) during which the upper and lower confidence bounds remain within the acceptance limit. The arrows in the previous listing give the maximum month at which the one-sided lower 95% confidence bound for each lot mean remains above an acceptance limit of 90.0%LC.

For the SICS and SISS models, the shelf life must be based on the worst-case batch. ICH Q1E specifies how much extrapolation is allowed. Generally the shelf life is limited to no more than two times the duration of the stability study (assuming that the 95% confidence bound remains within the acceptance specification for this period). In the example given here, shelf life is limited to 27 months based on projections for the worst-case Batch 8. When limited extrapolation of shelf life is accepted at time of approval, postapproval long-term data must be reported to FDA to verify the shelf life. Accelerated data are used to justify the temporary excursion of environmental conditions and support the limited extrapolation of long-term data.

22.6.4 Simple random batch case

If the goal of a stability analysis is to use the available stability data to predict stability performance for a *future* batch of product that (of course) was not studied, then the batch variable should be considered random. This should be reflected in the stability model and the analysis approach. A prerequisite for such a prediction is the availability of stability data from a reasonably large number of batches (hopefully more than three).

The appropriate predictive model is given by Eq. (22.1), where γ is a matrix of batch-specific random intercept and slope effect column vectors. The random intercept effect for a given batch may reflect random deviations from a target level of added active ingredients, whereas the random slope effect may reflect random deviations from nominal excipient levels or processing settings.

In this case, Eq. (22.1) simplifies to the following form:

$$y_{ij} = a_i + x_{ij}b_i + e_{ij} \tag{22.16}$$

where

y_{ij} is the observed test result for the ith batch tested at the jth time point
a_i and b_i are the random intercept and slope for the ith batch
x_{ij} is the storage time corresponding to the ith batch tested at the jth time point
$e_{ij} \sim iid\ N(0,\sigma^2)$ reflect measurement variance

If we assume that these batch-specific intercept and slope effects are normally distributed about process mean parameters α and β, respectively, and are mutually independent, then we take for the ith batch

$$\begin{pmatrix} a_i \\ b_i \end{pmatrix} \sim N\left(\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \begin{pmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_b^2 \end{pmatrix} \right) \tag{22.17}$$

where σ_a^2 and σ_b^2 represent the variances of these random intercept and slope effects among batches. This model may be analyzed using the following SAS MIXED procedure syntax:

```
proc mixed data=stability;
    class Batch;
    model Potency=Month /solution alphap=0.1
outp=outp;
    random Intercept Month/ type=vc
subject=Batch;
run;
```

When this code is used to analyze the eight batches of data in Appendix A, the parameter point estimates listed in Table 22.11 are obtained.

If a future lot is included in the incoming data set, whose test results are all missing, then the outp data set will contain the estimated one-sided lower 95% confidence bound for Potency at the storage times of interest. A partial listing of the outp data set is given here:

```
batch      Month      Lower
FUTURE      40      90.9427
FUTURE      41      90.7299
FUTURE      42      90.5169
FUTURE      43      90.3038
FUTURE      44      90.0904 <--
FUTURE      45      89.8769
```

TABLE 22.11 Point Estimates From Proc MIXED Random Batch Analysis

Parameter	Point estimate
A	101.46
β	− 0.2063
σ_a^2	1.4774
σ_b^2	0
σ^2	0.8063

TABLE 22.12 Output From SAS Fixed Batch Analysis

Covariance Parameter Estimates						
Cov Parm					Estimate	
Residual					1.1568	
Solution for Fixed Effects						
Standard						
Effect	batch	Estimate	Error	DF	t Value	Pr > |t|
Intercept		100.82	0.3840	24	262.52	< .0001
Month		−0.2131	0.02433	24	−8.76	< .0001
batch	BATCH3	1.3556	0.4849	24	2.80	0.0100
batch	BATCH4	3.4352	0.5032	24	6.83	< .0001
batch	BATCH5	0	.	.	.	.

```
FUTURE      46      89.6632
FUTURE      47      89.4493
FUTURE      48      89.2353
FUTURE      49      89.0210
```

The arrow in this code fragment indicates that the estimated shelf life for a future lot is 44 months.

22.6.5 Shelf-life estimation in more complex studies

The principles illustrated in the previous sections apply as well when the study includes multiple levels of strength, packaging, or other design variables. A good review of model building aspects in complex cases is provided by Millikin and Johnson.[32]

22.7 RELEASE LIMIT ESTIMATION

The shelf life estimation discussed previously depends on the true level at release (ie, the intercept) being in control. Thus, the true release levels of batches used for shelf life determination are assumed to be representative of those of future lots. Often, a process control limit is established to ensure good control of the release level. If the measured release level exceeds the limit, the product is not released for sale.

In keeping with the definition of shelf life, a one-sided release limit should ensure with at least 95% confidence that the mean level of a released lot will remain above (or below) a given lower (or upper) stability specification at the end of shelf life. Generally, stability data are available, and the objective is to calculate a release limit (RL) from the estimated slope, its standard error of estimate ($\hat{\sigma}_{slope}$), along with the established shelf-life acceptance limit (SL) of the desired shelf life (D), the estimated total analytical standard deviation ($\hat{\sigma}$), and the number of replicates averaged to obtain the reportable release test result (n). Then the release limit estimate is given by[33]

$$RL = SL - \delta \cdot D + t_{p,df}\sqrt{\hat{\sigma}^2/n + D^2 \cdot \hat{\sigma}^2_{slope}}, \qquad (22.18)$$

where

δ = min[$slope$,0] or max[$slope$,0] for lower or upper RL, respectively
$t_{p,df}$ = pth quantile of the t-distribution with df degrees of freedom.
df = Satterthwaite degrees of freedom determined as described in (Allen et al., 1991, page 1211)
p = 0.95 or 0.05 for lower or upper RL, respectively.

To illustrate this application of Eq. (22.18), the data from Appendix A will be used. A fixed batch analysis of batches 3, 4, and 5 yields a separate intercept, common slope model. The multiple regression output from a SAS analysis is shown in Table 22.12.

With a stability lower acceptance limit of SL = 90.0% LC and using the methods discussed here, shelf lives of 48, 57, and 43 months are obtained for batches 3, 4, and 5, respectively. As a result, a shelf life of 43 months could be justified for this product. Thus, the development team feels justified in recommending a shelf life for this product of D = 36 months; however, they also decide to establish a lower release limit on the mean of n = 2 replicate tests to ensure that no future batch mean will be below 90%LC through 36 months with 95% confidence. From the SAS output in Table 22.12, the following inputs to Eq. (22.18) are obtained:

$slope$ = −0.2131%LC/month

$\delta = \min[slope,0] = -0.2131$ (for lower release limit)
$\hat{\sigma} = 1.1568$ (residual error)
$\hat{\sigma}_{slope} = 0.02433$
$df = 24$
$p = 0.95$ (for lower release limit)

Application of Eq. (22.18) yields a release limit of $RL = 99.7\%$LC for this product.

For a separate intercept, separate slope model, a conservative approach is to take the slope estimate from the worst-case batch. In cases where the estimated slope implies a favorable divergence over time from the limit of interest (eg, negative slope used to estimate an upper RL), the slope should be set to zero, as implied by the δ function given previously. This provides a worst-case, conservative estimate for the release limit and ensures that lot mean levels will remain in conformance throughout the storage period with at least 95% confidence.

When estimates of $\hat{\sigma}$ and $\hat{\sigma}_{\text{slope}}$ are not obtained from the same regression, then the evaluation of df can be problematic. However, the Satterthwaite[34] approximation may be applied in this case. In principle, a similar approach could be used in the random batch case.

22.8 PROBABILITY OF FUTURE OOS STABILITY TEST RESULTS

The approaches described in the previous sections for estimating product shelf life are meant to control batch mean levels, but they give no assurance that individual test results, for lots placed on annual stability, will remain within the stability acceptance limit. Stability failures can result in costly investigations or, in the worst cases, product recall or withdrawal. Thus, after ensuring that all regulatory guidance has been followed, it is in a sponsor's interest to predict the probability of OOS stability test results for future batches.

The problem of prediction is fundamentally different than that of hypothesis testing and estimation and requires a full probability model of the process being simulated. Traditional data analysis methods have yielded ingenious analytical tools, such as tolerance and prediction interval estimates, as aids to risk assessment. However, except in the simplest scenarios, the prediction problem is too difficult to be solved analytically; modern computer simulation approaches are required. Methods such as mixed model regression rely on approximations that may not hold in small samples. Traditional tolerance and prediction intervals do not adequately address estimation uncertainty, and they rely on the nonintuitive paradigm of repeated sampling.

The prediction problem can be tackled in a very direct, rigorous, and informative way by taking a Bayesian perspective. Bayesian estimates are directly interpretable as probabilities. Modern tools for implementing the Bayesian approach, such as WinBUGS or the SAS MIXED procedure, are now readily available and can be effectively used to enhance decision making in the stability arena. Thus, this chapter would not be complete without an illustration of the power of these important methods by analyzing the data from all eight batches in Appendix A using Gibbs's sampling in WinBUGS.

The central objective of a Bayesian analysis is to estimate $p(\theta \mid \mathbf{Y})$, the joint posterior distribution of unknown model parameters given the data. From the joint posterior, the distribution of any function of θ or of any quantity that may be predicted from the model can be obtained through Markov chain Monte-Carlo techniques. By sampling iteratively from the posterior distribution, WinBUGS produces a posterior sample of the model parameters, θ from $p(\theta \mid \mathrm{Y})$.

A sample from the distribution of any predicted quantity (say, Z_{pred}) can be obtained by integrating the relevant component distributions in the data model over the posterior as follows:

$$p(Z_{\text{pred}}|\mathbf{Y}) = \int_{\theta} f(Z_{\text{pred}}|\theta)\cdot p(\theta|\mathbf{Y})d\theta$$

As with the posterior distribution, the distributions for predicted quantities are obtained within the process of Monte-Carlo iterative sampling and no analytical integration is needed. The example given in the next section is practical in nature and assumes only general familiarity with the Bayesian paradigm. For more depth and discussion of important topics such as convergence verification, readers are referred to a general text on Bayesian data analysis like Gelman et al.[35] For details of WinBUGS operation and syntax, see the WinBUGS operator's manual.[36]

22.8.1 Random batch model for prediction

Prediction involving future batches requires a random batch assumption so the model of Section 22.6.4 will be used with minor modification of symbols and structure. In place of Eq. (22.16), we define the true potency, μ_{bi}, for a given lot, b, tested at the ith time, X_i, as

$$\mu_{bi} = \alpha_b + \beta_b\cdot(X_i - Xbar),$$

where $Xbar$ is the average of the storage period for all nonmissing individual measurements in the data set. For the eight batches in Appendix A, $Xbar = 9.25$ months. This centering is used in the model to improve the rate of convergence of the Gibbs sampler.

Each of the $b = 1, 2, \ldots 8$ batches is assumed to have a unique true intercept, α_b, and true slope, β_b. The true slope and intercept for batches are assumed to be independent and to follow a normal distribution among lots, with the eight batches on hand representing a random sample from the population of all possible batches. Thus, as with Eq. (22.17), we have

$$\begin{pmatrix} \alpha_b \\ \beta_b \end{pmatrix} \sim N\left(\begin{pmatrix} \alpha_c \\ \beta_c \end{pmatrix}, \begin{pmatrix} \sigma_\alpha^2 & 0 \\ 0 & \sigma_\beta^2 \end{pmatrix} \right)$$

The observed potency result, Y_{bi}, for lot b at time point i will vary about the true value due to analytical variance and is assumed to follow a normal distribution as follows:

$$Y_{bi} \sim N(\mu_{bi}, \sigma^2).$$

22.8.2 Prior distributions for model parameters

The probability model described previously depends on a vector of five model parameters:

$$\boldsymbol{\theta} = [\alpha_c, \beta_c, \sigma_\alpha, \sigma_\beta, \sigma]$$

A Bayesian perspective considers these parameters to be random quantities and thus to have a probability distribution. Prior to data collection, the distributions of these parameters are a quantitative expression of the subjective uncertainty about their value. When prior uncertainty is great, it is appropriate to assume prior distributions that are independent and very broad or noninformative. This was the approach taken here. The following noninformative distributions were used:

$$\begin{aligned} &\alpha_c \sim N(100, 10^{-6}) \\ &\beta_c \sim N(0, 10^{-6}) \\ &\sigma_\alpha = 1/\sqrt{\tau_\alpha} \sim U(0, 100) \\ &\sigma_\beta = 1/\sqrt{\tau_\beta} \sim U(0, 100) \\ &\tau_c = 1/\sigma^2 \sim G(10^{-3}, 10^{-3}) \end{aligned}$$

U(A,B) represents a flat uniform distribution with minimum and maximum values A and B, respectively. G(C,D) represents a gamma distribution with shape and rate parameters C and D, respectively. The parameters for all these distributions are chosen so that the probability density is essentially constant over the likely range of the true parameter values; thus, the choice of prior has essentially little or no impact on the final estimation. In cases were substantial prior knowledge and theory on a product, or like products, are available, it may be appropriate to use a more informative prior, but that will not be discussed here.

22.8.3 Predicted quantities of interest

The objective of the present analysis is to simulate probability distributions of various predicted quantities that are functions of the original parameters. These consist of posterior or posterior predictive distributions that summarize the available knowledge about each predicted quantity.

One quantity of interest is the initial mean potency for the overall manufacturing process, Y_0, where

$$Y_0 = \alpha_c - \beta_c \cdot Xbar$$

Ideally, Y_0, should be close to 100%LC. If this value is significantly below 100%LC, it may mean that the active ingredient is being lost in the manufacturing process, and an increase in overage is needed to correct this situation.

Certainly, it would also be of interest to estimate the distributions for a future batch slope, β_{pred}, and initial potency level, $Init_{\text{pred}}$, which can be obtained from the model parameters as follows:

$$\begin{aligned} &\beta_{\text{pred}} \sim N(\beta_c, \sigma_\beta^2) \\ &Init_{\text{pred}} = \alpha_{\text{pred}} - \beta_{\text{pred}} \cdot Xbar, \quad \text{where} \quad \alpha_{\text{pred}} \sim N(\alpha_c, \sigma_\alpha^2) \end{aligned}$$

From these distributions, one can also predict the true mean level, μ_{pred}, and individual potency test result, Y_{pred}, of some future lot after X_{pred} months of storage (for this example, we take $X_{\text{pred}} = 36$ months). This can be obtained as described in the following sections.

Also of particular interest would be the posterior probability distribution of analytical error, σ, and the predictive posterior distribution of the product shelf life. The shelf life for a given batch will be the storage period during which the true level remains above the stability acceptance limit SL:

$$Shelf.Life = \frac{SL - Init_{\text{pred}}}{\beta_{\text{pred}}}$$

For this example, we take $SL = 90\%$LC. Some care is needed in the definition of shelf life. Since the distribution of β_{pred} may include nonnegative slope values, which could result in inadmissible (negative or infinite) values of *Shelf.Life*. In the WinBUGS code given next, the *Shelf.Life* values are truncated in the interval at 0 and 100 months.

22.8.4 Implementation in WinBUGS

This application is easily programmed into the freely available WinBUGS software.[36] The WinBUGS code consists of the following three text sections that are placed in a WinBUGS file: Model, Data, and Initial values (to start the Markov chain Monte Carlo iterations). These are discussed briefly in the next sections.

22.8.4.1 Model

The Model code consists of the following text:

```
model;
{ for(batch in 1 : B) {

              alpha[batch]   ~   dnorm(alpha.c,tau.
alpha)
              beta[batch] ~ dnorm(beta.c,tau.beta)
              for(i in 1 : N) {
                mu[batch, i] <- alpha[batch] + beta
[batch] * (X[i] - Xbar)
                Y[batch, i] ~ dnorm(mu[batch, i],
tau.c)
              }
      }
          # Priors
            tau.c ~ dgamma(0.001,0.001)
            beta.c ~ dnorm(0.0,1.0E-6)
            alpha.c ~ dnorm(100,1.0E-6)
            sigma.beta ~ dunif(0,100)
            sigma.alpha ~ dunif(0,100)
            tau.beta <- 1 / (sigma.beta * sigma.beta)
            tau.alpha <- 1 / (sigma.alpha * sigma.
alpha)
            sigma <- 1 / sqrt(tau.c) # Total analyti-
cal SD
          # Predicted Quantities of Interest
              # Process Initial Target
              Y0 <- alpha.c - beta.c*Xbar
              beta.pred ~ dnorm(beta.c,tau.beta) #
Lot Slopes
              alpha.pred ~ dnorm(alpha.c,tau.alpha)
              # Lot Initials
              Init.pred <-alpha.pred - Xbar*beta.
pred
              #Shelf life constrained to [0,100]
              Shelf.Life<-max(0,min(100,(90-Init.
pred)/beta.pred))
              # Predicted Lot Means at Xpred
              mupred <-alpha.pred + (Xpred-Xbar) *
beta.pred
              # Individual Prediction at Xbar
              Ypred ~ dnorm(mupred,tau.c)
}
```

22.8.4.2 Data

The potency stability data used for analysis is given in Appendix A. These consist of data from eight batches. Measurements were taken at up to 11 nominal time points over a period of 24 months. The WinBUGS code for these data is given here:

```
# Data
list(X = c(0, 1, 2, 3, 3, 6, 6, 12, 12, 24, 24),
B = 8,N = 11,Xpred = 36,Xbar = 9.25,
Y = structure(.Data = c(102,NA,NA,101,
101,100.5,100.8,
100.1,99,96.7,97.2,
101,101.3,NA,99.8,99.2,99.5,97.8,97.4,97.2,
96.9,96,
104.8,NA,NA,103,101.2,100.8,99.2,98.6,97.2,
97.6,98,
104,NA,NA,103.2,NA,102.8,103.3,102.4,101.2,
99.1,99.5,
102,101.4,100.8,100.2,99.7,98.8,98.5,98,97.1,
96.6,96.1,
102.7,102,NA,100.6,99.9,99.2,98.5,98.1,97.6,
96.8,96,
101.3,101.5,NA,100.2,99.8,99,98.5,98.5,97.4,
96.6,96.4,
101.6,NA,NA,100,NA,99,NA,97.8,97,NA,NA),.
Dim = c(8,11)))
```

22.8.4.3 Initial values

Gibbs sampling uses Monte Carlo sampling from the various prior, model, and predictive distributions indicated previously. The sampling is dependent (not pseudorandom) because the sampling at any iteration depends on the values in the previous iteration; however, the sampling procedure is known to converge on the desired posterior distribution. Thus, starting values are needed to initiate the Gibbs sampling process. Examination of the iterative process showed that convergence to the final distribution was rapid and that the process did not depend on the initial starting values. However, the results reported here were obtained using the following set of starting values (the 0 superscript indicates a starting value for initiating Markov chain–Monte Carlo sampling).

Model parameters:

$$\alpha_c^0 = 100,\ \ \beta_c^0 = 0,\ \ \sigma_\alpha^0 = 1,\ \ \sigma_\beta^0 = 1,\ \ \tau_c^0 = 1$$

Latent quantities for the eight lots in the data set:

$$\alpha_b^0 = 100,\ \ \beta_b^0 = 0,\ \ b = 1, 2, \ldots 18$$

Predicted quantities:

$$\alpha_{\text{pred}}^0 = 100,\ \ \beta_{\text{pred}}^0 = 0,\ \ Y_{\text{pred}}^0 = 100$$

Missing Potency values in the data set:

$$Y_{bi}^0 = 100\ \ \forall\ \ Y_{bi} = NA$$

These initial values were specified by the following WinBUGS code:

```
# Initials
list(tau.c = 1,sigma.beta = 1,beta.pred = 0,beta.
c = 0,sigma.alpha = 1,

      alpha.pred = 100,alpha.c = 100,Ypred = 100,
        beta = c(0,0,0,0,0,0,0,0),
      alpha = c(100,100,100,100,100,100,100,100),

Y = structure(.Data = c(NA,100,100,NA,NA,NA,NA,
NA,NA,NA,NA,
```

```
NA,NA,100,NA,NA,NA,NA,NA,NA,NA,NA,
NA,100,100,NA,NA,NA,NA,NA,NA,NA,NA,
NA,100,100,NA,100,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,100,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,100,NA,NA,NA,NA,NA,NA,NA,NA,
NA,100,100,NA,100,NA,100,NA,NA,100,100),.Dim=c
(8,11))
```

22.8.4.4 *Burn-in period, sample size, and convergence verification*

A *burn-in period* is recommended for Markov chain–Monte Carlo sampling to give the dependent draws a chance to migrate from the initial starting values and approximate the target distribution. A close examination of the sampling process in WinBUGS showed it was very well behaved and rapidly converged. However a burn-in period of 5000 draws was still used before the samples were used to estimate the distributions of interest. A subsequent sample of 10,000 draws was used for estimation. This large sample size ensured that the Monte Carlo sampling error was insignificant compared to the variance of the target distributions of all quantities estimated.

The 1000-draw predictive posterior sample was exported from WinBUGS (coda option) and imported as a text file into JMP.[37] Kernel density histograms of the posteriors are provided in Fig. 22.3. The OOS probability was calculated as a percentage of the predictive posterior samples below 95%LC (given in Table 22.13).

22.8.5 Results

To obtain more specific information on any posterior or predictive posterior sample, it may be exported from WinBUGS and imported into a more general data analysis package such as SAS or JMP to conduct capability analyses or obtain more specific quantiles. The histogram obtained for *Shelf.Life* is shown in Fig. 22.4.

The shaded region in the figure identifies the 5% lowest *Shelf.Life* values. The fifth percentile of this distribution was 40.3 months. Thus, 95% of future batches are expected to have a *Shelf.Life* greater than 40 months. Therefore, a shelf-life estimate of about 40 months would be consistent with the FDA definition

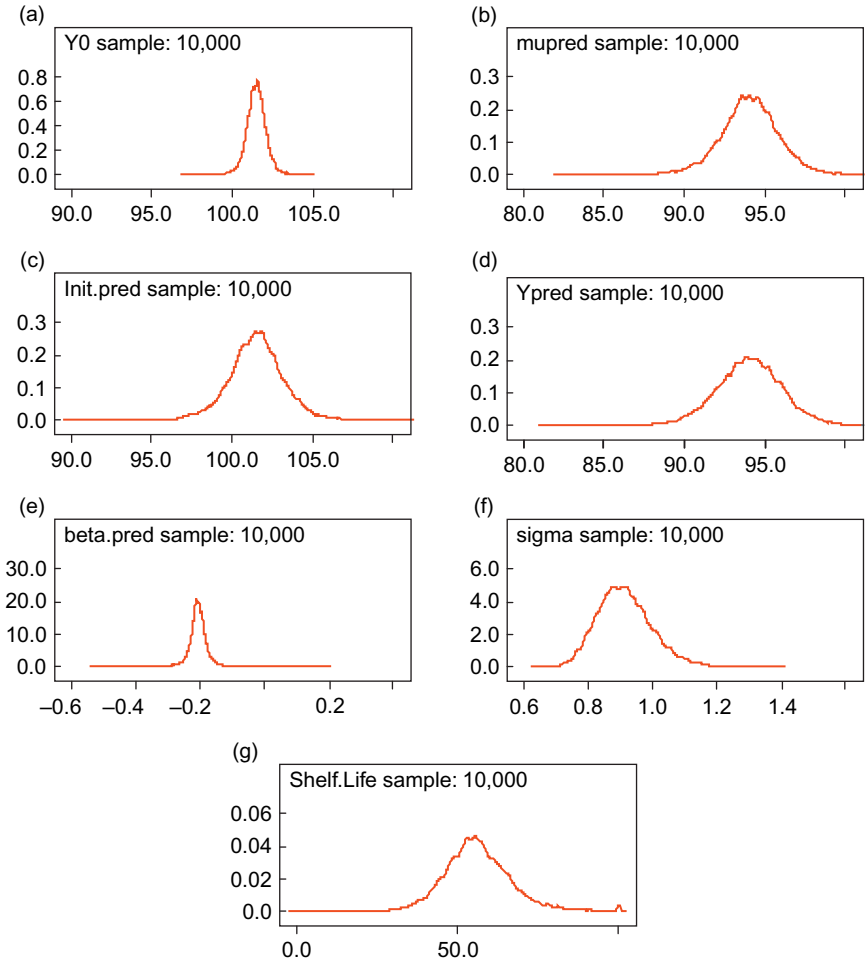


FIGURE 22.3 Kernel density smoothed posterior or predictive posterior distributions of quantities of interest. (a) Manufacturing process true initial mean. (b) True future batch potency at 36 months. (c) True initial means for future batches. (d) Future individual test results at 36 months. (e) True future batch slope. (f) Analytical standard deviation. (g) Shelf life for future batches.

TABLE 22.13 Posterior and Predictive Posterior Distribution Statistics for Quantities of Interest

Predicted quantity	Distribution mean	Distribution standard deviation	2.5% Percentile	Median (50% percentile)	97.5% Percentile
$Init_{pred}$	101.5	1.782	97.85	101.5	105.0
Shelf.Life	56.28	11.08	36.32	55.57	81.28
Y_0	101.5	0.6081	100.2	101.5	102.7
Y_{pred}	94.01	2.135	89.81	94.03	98.19
β_{pred}	− 0.2068	0.02959	− 0.2662	− 0.2071	− 0.148
μ_{pred}	94.03	1.922	90.1	94.04	97.79
σ	0.9139	0.08429	0.7675	0.9078	1.099

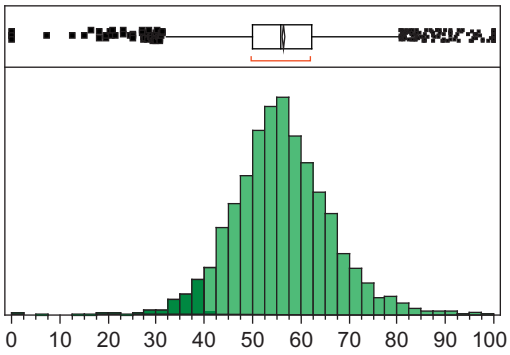


FIGURE 22.4 Predictive posterior distribution of shelf life.

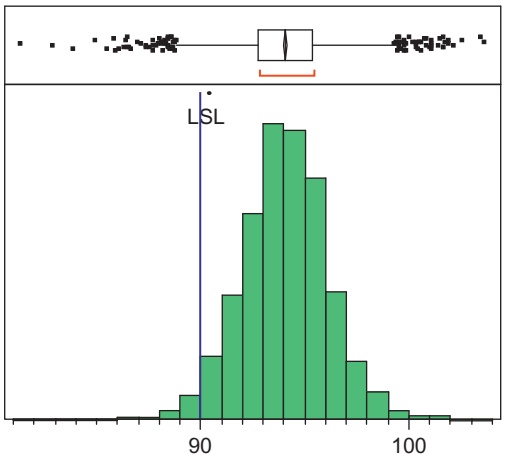


FIGURE 22.5 Process capability of predictive posterior of individual test results at 36 months.

that ensures that batch means will remain within acceptance limits with 95% confidence.

The process capability analysis conducted on Y_{pred} is shown in Fig. 22.5.

The area to the left of the acceptance limit of 90%LC contains 2.8% of the predictive posterior distribution. Thus, future lots can expect to have OOS test results at the 36-month time point with a probability of 0.028. In view of this risk, an appropriately shorter shelf-life or release limit could be considered to avoid the cost of investigations or (in the worst case) the disruption of a product recall.

In stability investigations, predictions of future stability test results for a particular batch of interest might be of interest. In the data set given in Appendix A, future values are missing (*NA*). In WinBUGS, missing values are treated as predictive quantities to be estimated. Thus, we may directly request predictive posteriors for future (missing) data values that have not yet been obtained.

22.8.6 Bayesian prediction using SAS proc MIXED

The SAS MIXED procedure can provide a posterior sample from a mixed model fit, such as the random batch model used with WinBUGS:

```
proc mixed;
   class Batch;
   model Potency = Month /solution alphap = 0.1
outp = outp;
   random Intercept Month/ type = vc subject = Batch;
   prior /out = posterior nsample = 10000;
run;
```

By default, noninformative priors are used for both fixed and random parameters. In this sample code, the posterior data set will contain 10,000 draws from the posterior. This data set may be used to further generate samples of parameter transformations, including predictive posterior for future results, as was illustrated in WinBUGS in the previous section.

TABLE A.1 Sample Data From Shuirmann[38]

Number of months	Batch 1	Batch 2	Batch 3	Batch 4	Batch 5	Batch 6	Batch 7	Batch 8
0	102	101	104.8	104	102	102.7	101.3	101.6
1	–	101.3	–	–	101.4	102	101.5	–
2	–	–	–	–	100.8	–	–	–
3	101	99.8	103	103.2	100.2	100.6	100.2	100
3	101	99.2	101.2	–	99.7	99.9	99.8	–
6	100.5	99.5	100.8	102.8	98.8	99.2	99	99
6	100.8	97.8	99.2	103.3	98.5	98.5	98.5	–
12	100.1	97.4	98.6	102.4	98	98.1	98.5	97.8
12	99	97.2	97.2	101.2	97.1	97.6	97.4	97
24	96.7	96.9	97.6	99.1	96.6	96.8	96.6	–
24	97.2	96	98	99.5	96.1	96	96.4	–

– indicates test not available.

APPENDIX A SAMPLE DATA

The following stability data set is used as an example in this chapter. It is taken from Shuirmann[38] (Table A.1).

SAS code for reading these data into a data set for analysis is given here:

```
data stability;

  input Month@;
  do batch = 'BATCH1','BATCH2','BATCH3','BATCH4',
             'BATCH5','BATCH6','BATCH7','BATCH8';
        input Potency@;output;
  end;
cards;

0  102   101   104.8 104   102   102.7 101.3 101.6
1  .     101.3 .     .     101.4 102   101.5 .
2  .     .     .     .     100.8 .     .     .
3  101   99.8  103   103.2 100.2 100.6 100.2 100
3  101   99.2  101.2 .     99.7  99.9  99.8  .
6  100.5 99.5  100.8 102.8 98.8  99.2  99    99
6  100.8 97.8  99.2  103.3 98.5  98.5  98.5  .
12 100.1 97.4  98.6  102.4 98    98.1  98.5  97.8
12 99    97.2  97.2  101.2 97.1  97.6  97.4  97
24 96.7  96.9  97.6  99.1  96.6  96.8  96.6  .
24 97.2  96    98    99.5  96.1  96    96.4  .

                                        run;
```

References

1. Q1A(R2) Stability testing of new drug substances and products; November 2003.
2. Q1B Photostability testing of new drug substances and products; November 1996.
3. Q1C Stability testing for new dosage forms; May 1997.
4. Q1D Bracketing and matrixing designs for stability testing of new drug substances and products; January 2003.
5. Q1E Evaluation of stability data; June 2004.
6. Q2A Text on validation of analytical procedures; October, 1994.
7. Q2B Validation of analytical procedures: methodology; November 1996.
8. Q3A Impurities in new drug substances; February 2003.
9. Q3B(R) Impurities in new drug products; February 2003.
10. Q6A Specifications: test procedures and acceptance criteria for new drug substances and new drug products: chemical substances; December 2000.
11. Q6B Specifications: test procedures and acceptance criteria for biotechnological/biological products; August 1999.
12. Tsong Y. Recent issues in stability study. *J Biopharm Stat* 2003;**13**(3):vii–ix.
13. The SAS system for statistical analysis is available from the SAS institute, Inc., Cary, NC.
14. Neter J, Kutner MH, Nachtsheim CJ, Wasserman W. *Applied linear statistical models*. 4th ed. Chicago, IL: Irwin/McGraw-Hill, Inc.; 1996.
15. Carstensen JT. *Drug stability: principles and practices*. New York, NY: Marcel Decker, Inc.; 1995ISBM0-8247-9635-7, Chapters 2–10
16. Lin TYD, Chen CW. Overview of stability study design. *J Biopharm Stat* 2003;**13**(3):337–54.
17. Box GEP, Hunter WG, Hunter JS. *Statistics for experimenters*. NY: John Wiley & Sons, Inc; 1978.
18. Nordbrock E. Design and analysis of stability studies. In: ASA proceedings of the biopharmaceutical section; 1994, p. 291–294.
19. Nordbrock E. Statistical comparison of stability study designs. *J Biopharm Stat* 1992;**2**(1):91–113.
20. The R package and associated operators manual are available for download from the following website: http://www.r-project.org/.

21. Kiermeier A, Jarrett RG, Vebyla AP. A new approach to estimating shelf-life. *Pharm Stat* 2004;**3**:3–11.
22. Chow SC, Liu JP. *Statistical design and analysis in pharmaceutical science*. New York, NY: Marcel Dekker, Inc; 1995.
23. Shao J, Chow SC. Statistical inference in stability analysis. *Biometrics* 1995;**50**:753–63.
24. Fairweather WR, Lin TYD, Kelly R. Regulatory, design, and analysis aspects of complex stability studies. *J Pharm Sci* 1995;**84**(11):1322–6.
25. Chen JJ, Ahn H, Tsong Y. Shelf life estimation for multifactor stability studies. *Drug Inf J* 1997;**31**:573–87.
26. Chen JJ, Hwang JS, Tsong Y. Estimation of the shelf-life of drugs with mixed effects models. *J Biopharm Stat* 1995;**5**(1):131–40.
27. Chow SC, Shao J. Estimating drug shelf-life with random batches. *Biometrics* 1991;**47**:1071–9.
28. Ng MJ. STAB stability system for SAS, Division of Biometrics, CDER, FDA. These SAS macros may be downloaded from the following website: http://www.fda.gov/cder/sas/; 1995.
29. Brownlee KA. *Statistical theory and methodology*. Malabar, FL: Robert E. Keiger; 1984.
30. Bancroft TA. Analysis and inference for incompletely specified models involving the use of preliminary test(s) of significance. *Biometrics* 1964;**20**:427–42.
31. Ruberg SJ, Stegman JW. Pooling data for stability studies: testing the equality of batch degradation slopes. *Biometrics* 1991;**47**:1059–69.
32. Millikin GA, Johnson DE. *Analysis of messy data*, vol. III. ANCOVA. Boca Raton, FL: Chapman and Hall and CRC Press 1992.
33. Allen PV, Dukes GR, Gerger ME. Determination of release limits: a general methodology. *Pharm Res* 1991;**8**(9):1210–13.
34. Satterthwaitte FE. An approximate distribution of estimates of variance components. *Biometrics Bull* 1946;**2**:110–14.
35. Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian data analysis*. Boca Raton, FL: Chapman & Hall/CRC Press; 2004.
36. The WinBUGS package was developed jointly by the Cambridge MRC and the Biostatistics Unit at the Imperial College School of Medicine, London. The software and operators manuals may be downloaded from the following website: http://www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml.
37. Reference to the SAS JMP system.
38. Shuirmann DJ. Current statistical approaches in the Center for Drug Evaluation and Research.In: FDA, Procedings of Stability Guidelines, AAPS and FDA Joint Conference, Arlington, VA, December 11–12; 1989.

This page intentionally left blank

CHAPTER

23

Packaging Selection for Solid Oral Dosage Forms

Y. Chen

Novast Laboratories, Ltd., Nantong, China

23.1 INTRODUCTION

Drug products are always packaged in containers for storage and marketing for convenience of handling, use, and product protection. Stability information of a drug product in selected containers regarding the proposed shelf life is part of the quality information required to obtain approval from regulatory agencies to market any drug product, in addition to proof that the product is safe, efficacious, high quality, and made in compliance with current good manufacturing practices. Product stability is typically demonstrated through real-life, formal stability studies under ICH conditions.[1] Study designs and data analysis of long-term studies under the International Conference on Harmonization (ICH) conditions have been discussed in chapter "Statistical Design and Analysis of Long-Term Stability Studies for Drug Products." Due to the time-consuming nature of long-term studies, it is very important to select suitable containers to ensure the success of the studies, particularly when packaging protection of products is needed. Product stability may be imparted by formulation design and proper use of containers. Formulation design is most critical in developing stable drug products. However, formulation design itself may not always be sufficient to ensure that the product is stable during storage. Due to the vast diversity of drug properties and their interaction with environmental conditions, some drug products require protection against some degrading factors, such as humidity and oxygen. The selection process to identify the proper packaging for the needed protection is, therefore, an important part of product development. Although one could argue that the selection could be very simple by using the most highly protective, or the highest barrier container-closure systems for any product, the high cost for the use of such high barrier container closure systems may be difficult to justify, or impractical for commercialization unless the high barrier container closure system is needed for protection of the product. Therefore, one must select appropriate container closure systems to balance the need for product protection and the cost for commercialization. Formal stability study is the definite process to demonstrate the quality of any drug product during its shelf life in the selected container closure system. However, real-life stability study is not the best process to select containers because the process is empirical, time consuming, costly, and inefficient, making the study unsuitable for screening purposes. Any failure of a formal stability study could significantly delay the time-to-market for the product, diminish the opportunity for providing disease treatment to patients, and lead to significant financial losses for a drug maker. Instead, proper selection of containers can be achieved prior to the formal stability study, using science-based and quality-by-design (QbD) approaches throughout early development stages. Containers can be selected based on material characteristics, and by linking the properties of drug products with that of the containers. Formal stability studies can then be conducted to prove that the product is stable in the selected container closure systems. This chapter will focus on the science-based selection of containers and/or packaging materials for protection of solid oral dosage (SOD) forms.

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00023-6

23.1.1 Definitions

To avoid any possible confusion, several terms related to packaging and containers are defined according to FDA guidance.[2,3]

- A *packaging component* means any single part of a container closure system.
- A *container closure system* refers to the sum of packaging components that together contain and protect the dosage form. This includes primary packaging components and secondary packaging components (if the latter are intended to provide additional protection to the drug product).
- A *packaging system* is equivalent to a container closure system.

  In this chapter, container, container closure system, and packaging are interchangeable.
- *Primary packaging component* refers to a packaging component that is or may be in direct contact with the dosage form (eg, liners, bottles, desiccant containers in bottles with dosage forms, blister films, and lidding foils).
- *Secondary packaging component* means a packaging component that is not, and will not be, in direct contact with the dosage form (eg, cartons, and over-wraps for blister cards).

23.1.2 General considerations

Packaging materials used for pharmaceutical products must meet regulatory[2,3] and compendial requirements,[4] provide adequate protection of products against degrading and/or contaminating factors, function properly as intended, and be compatible with drug products and safe for use. In addition, packaging should be elegant for marketing, convenient for use, and low cost for the benefit of both patients and product manufacturers.

Containers are used for different types of products for different routes of administration. By routes of administration, pharmaceutical products can be classified as inhalation, injectable, ophthalmic, liquid oral, liquid topical, transdermal, suppository, and SOD forms. Container requirements for different types of products may be different. Concerns for container-drug interaction are the highest for injectable and inhalation products, and lowest for SOD forms. Therefore, a container closure system approved for one type of product cannot be assumed to be approved for another type of product. For detailed information on regulatory requirements for different types of drug products, refer to the FDA guidance *Container Closure Systems for Packaging Human Drugs and Biologics*.[2]

Proper use of container closure systems can provide protection of SOD forms from moisture, light, oxygen or other gases, microbial contamination, and mechanical stress, depending on the protection requirements of the product. In general, protection against moisture, light, and oxygen are the most important factors to consider. Resistance to microbial contamination can be readily achieved using well sealed containers with proper control of materials, manufacturing, and packaging processes under good manufacturing practice conditions. Protection against physical stress can be achieved by the use of containers with sufficient strength and minimized headspace, which prevents excessive movement of products during handling and shipping. Containers for finished products are normally incapable of protecting products from thermal stress. Heat-sensitive products are typically protected by refrigeration or freezing during storage, shipping, and end use. Therefore, protection against microbial contamination, mechanical stress, and thermal stress, will not be discussed in detail in this chapter.

As stated previously, formal stability studies are not desirable for screening containers for pharmaceutical products. Instead, containers should be selected based on scientific knowledge of the fundamental properties of containers and drug products. Rational selection of containers will minimize the risk of failure in stability studies, increase the efficiency of product development, and reduce the cost for both the pharmaceutical industry and patients.

23.2 MATERIAL CONSIDERATIONS

23.2.1 Containers

Containers can be classified by the dosage forms for which they are intended, by product contact, by container properties, or by purpose of use. Classifications based on dosage forms, such as injectable or SOD forms, and on product contact, such as primary or secondary packaging, were provided in the introduction section. Alternatively, containers can be classified based on their gas permeability—either as permeable or impermeable containers. In general, polymeric blisters and bottles, such as polyvinyl chloride (PVC) and polyvinylidene chloride (PVDC) blisters, and polyethylene terephthalate (PET) and high-density polyethylene (HDPE) or polypropylene (PP) bottles, are permeable to moisture and oxygen; while properly sealed glass bottles and aluminum foil/foil blisters can be considered impermeable to gases if there are no imperfections in the bottle seal or pinholes in the foil. Containers are also commonly classified as bulk containers and finished containers for use purposes.

Bulk containers are used for storage of bulk drug products prior to packaging or for shipping of bulk products to repackagers or contract packagers. Finished containers are for the final packages used to market and deliver drug products to patients. Finished containers are further classified into single-unit containers, and multiple-unit containers. Pharmacy containers for dispensing drug products to patients can be classified as finished containers. Stability data of products stored in the finished containers under ICH conditions are the main source of data required by regulatory agencies for approval of shelf life for products.

Bulk Containers: For bulk products of SOD forms, fiber and plastic drums with flexible liners, such as low-density polyethylene (LDPE, or PE) liners, are commonly used. The liners provide a controlled material contact surface for the products, and some protection of products against moisture and contamination. Regular fiber drums and PE liners are highly permeable to moisture and other gases. A fiber drum reinforced with aluminum foil may have significantly lower gas permeability than a regular one. Plastic drums or pails, such as HDPE drums, also have low gas permeability. With the use of heat-sealable laminated foil bags as liners, HDPE drums or fiber drums reinforced with aluminum foil are likely the second-best bulk containers for products that are sensitive to moisture or oxygen. Well-sealed stainless steel containers provide the best barrier for any gases. However, large steel containers are heavy, and thus inconvenient to use.

Finished containers: For finished pharmaceutical dosage forms, plastic bottles and flexible packages, such as blisters and pouches, are the most commonly used containers. However, plastic bottles, such as LDPE, HDPE, PET, polyethylene terephthalate G (PETG), and PP bottles, and polymeric blisters, such as PVC, PVDC, and polychloro-trifluoroethylene (PCTFE), are permeable to moisture and gases, although the permeability varies with materials. Glass bottles are not commonly used due to the safety concern for broken glasses during handling, and due to their heavy weight that can increase shipment cost, even though well-sealed glass bottles can be considered impermeable to gases. Foil/foil blisters or pouches in the absence of pinholes are impermeable to gases, but are not commonly used in the United States, due to the lack of transparency of the materials, and the low resistance to physical stress such that the blisters are easily deformed by external pressure or force. Foil/foil blisters are widely used in some countries in Asia for marketing reasons other than the need of product protection.

Container permeability is a critical factor to consider in selecting packaging materials for products that are sensitive to the respective permeants. Moisture and oxygen permeabilities of packaging materials are commonly measured, and reported as moisture/water vapor transmission rate (MVTR, or WVTR) and oxygen transmission rate (OTR), respectively. Table 23.1 lists examples of MVTR and OTR data from the literature for some packaging materials. It is important to note that the data listed in the table for blistering materials are for flat films. Permeation rate of the formed cavity of blisters will be larger than that of the flat films due to reduced thickness of the formed cavity. The actual dimensions, including surface area and thickness, of any given containers shall be accounted for in packaging design and evaluation. Increasing the thickness of the barrier material can reduce the MVTR of both the laminated flat films and the formed cavities, as shown for commercially available laminated film materials Aclar RX 160, SupRx 900, UltRx 2000, UltRx 3000, UltRx 4000, and UltRx 6000 (Table 23.2) if the dimension of the blister cavity is the same for different films with different thickness.

TABLE 23.1 Moisture and Oxygen Transmission Rates of Some Packaging Materials

Material	MVTR (g.mil/100 in^2 per 24 h, at 100°F, 90%RH)	OTR (cc.mil/100 in^2 per 24 h-atm, at 77°F)	Reference
Cold form foil	0.00	0.00	6
Polychloro-trifluoroethylene (PCTFE)	0.016	7.00	6
Polyvinyl dichloride (PVDC)	0.1–0.2[a]	0.15–0.90	7
High density polyethylene (HDPE)	0.3–0.4	139–150	6,8
Polypropylene (PP)	0.69–1.0	182	7,9
Low density polyethylene (LDPE)	1.0–1.5	463–500	6
Polyethylene terephthalate (PET)	1.2–2.0	3–5	7
PET modified with CHDM (PETG)	4	6.6	10
Polyvinyl chloride (PVC)	0.9–5.1	5–20	11
Ethylene vinyl alcohol (EVOH)	1.4–5.4[a]	0.05–0.90	7
Polystyrene (PS)	7–10	350–400	11
Nylon	16–20	1.0	11

[a]*Determined at 104°F/90%RH.*

TABLE 23.2 Barrier Thickness of Aclar Film and the Resulting MVTR

Aclar film	Aclar thickness (μm)	MVTR (mg/cm^2 per day at 37.8°C, 100%RH)PPP[a]
Rx 160	15	0.0419
Rx 20E	20	0.0295
SupRx 900	23	0.0264
UltRx 2000	51	0.0119
UltRx 3000	76	0.0088
UltRx 4000	102	0.0065
UltRx 6000	153	0.0038

[a]*Data from https://www.honeywell-aclar.com/ProductInformation/Pages/Documentation.aspx*

The data in Table 23.1 show that materials that have a low MVTR may not have a low OTR. For example, the MVTR of nylon is approximately 1000 times larger than that of PCTFE. However, the OTR of nylon is actually smaller than the OTR of PCTFE. For the purpose of moisture protection, PCTFE is far superior to all other materials listed in the table, except for cold form foil. On the other hand, PVDC and ethylene vinyl alcohol (EVOH) are better choices for protection against oxygen. Packaging materials should therefore be selected based on the specific need for protection of each individual product. From a material development point of view, data listed in Table 23.1 show that combination of PCTFE with EVOH can produce a multiple-layer blistering material with excellent barrier properties against moisture and oxygen. In fact, commercial products of such composite laminated material, PVC/EVOH/Aclar, have become available on the market in recent years.

Although aluminum foil may be considered as the ultimate barrier for gases, thin foil (thickness <25 μm) may have pinholes,[5] and the foil may become permeable through the pinholes. Pinholes can be caused either by the presence of organic contaminates in the molten aluminum or by the stress of the rolling process during foil production. The number of pinholes per defined area of the foil increase as the thickness of the foil decreases. This property is important to consider in selecting foil materials, and also in the design of packages using thin foil as a barrier.

23.2.2 Determination of container MVTR

As previously discussed, container moisture permeability is an important factor to consider in packaging selection for moisture-sensitive products. Reliable MVTR is critical to ensure the selection of proper materials. Compendial methods for determining MVTR are currently described in the USP Chapter <671>[4] to evaluate the performance of containers. Since the 37th edition, the USP Chapter <671> provides different MVTR test methods for multiple-unit containers for SOD forms. These methods are intended for either barrier protection determination, or classification of containers. It is important to use the right method to obtain reliable MVTR. For container classification purpose, the USP method uses a single weight gain measurement over a 14-day study through the use of control bottles. Data in literature have shown that results obtained by this method may be highly variable. For example, the standard deviation of the single weight gain results was higher than the mean value[12] in some cases. In one case, the value obtained at 25°C/75%RH was unreasonably higher than that measured at 40°C/75%RH.[12] A high variability in permeability renders the data of little scientific value, either for packaging design or for justification of container postapproval changes. Therefore, this single weight gain method is not suitable for determining container barrier protection property.

For barrier protection determination, the current version of USP Chapter <671> provides a steady state method (Method 1) with multiple time-point measurements for multiple-unit containers, and for high barrier single-unit containers (Method 2), to ensure the MVTR data is reliable and specific for discriminating the barrier performance of containers. The theory that supports the use of the steady state method for MVTR measurement is discussed next.

For well-sealed plastic containers without gross defects, the mechanism of water ingression into a container is mainly by diffusion of moisture molecules through the container wall,[13] and through the seal. For a relatively large container with a thin wall, the wall can be approximated as a homogeneous planar barrier. If it is assumed that:

1. the diffusion coefficient D of water vapor in the wall is constant at constant temperature;
2. the concentration of moisture at the outside surface is maintained at a constant level of C_o; and
3. the concentration at the inside surface is kept at essentially zero or a "sink" condition; then the initial and boundary conditions for the packaging system under testing can be described as:

$$C = C_i,\ 0 < x < h,\ t = 0$$
$$C = C_o,\ x = 0;\ C = 0,\ x = h,\ t \geq 0$$

where:

x is the direction of diffusion
h represents the thickness of the wall

C is the moisture concentration in the wall
C_i indicates the initial moisture concentration in the wall
C_o is the concentration of moisture at the outside surface of the container wall
t is time.

If the quantity of water diffused through the wall and entered in the "sink" is measured as a result of net diffusion, Eq. (23.1) can be used to describe the permeation process.[14]

$$Q_{t,d}=\frac{DC_o}{h}\left(t+\frac{C_ih^2}{2DC_o}-\frac{h^2}{6D}\right)-\frac{2hC_o}{\pi^2}\sum_{n=1}^{n=\infty}\frac{(-1)^n}{n^2}\exp\left(-\frac{n^2\pi^2Dt}{h^2}\right)$$
$$-\frac{4C_ih}{\pi^2}\sum_{m=0}^{m=\infty}\frac{1}{(2m+1)^2}\exp\left(\frac{-(2m+1)^2\pi^2Dt}{h^2}\right) \quad (23.1)$$

where:
$Q_{t,d}$ is the amount of moisture diffused through unit surface area of the wall at time t.

If the amount of water entering the outer surface of the wall material is measured, such as by measuring total weight gain (wg) of containers per USP Chapter <671>, the measured quantity per unit surface area by weight gain, $Q_{t,wg}$, includes the amount of water absorbed by the wall and the backing material in the closure, and the amount of water diffused through the wall into a "sink" created by desiccants. In this case, Eq. (23.2) can be used to describe the permeation process:

$$Q_{t,wg}=\frac{DC_o}{h}\left(t-\frac{C_ih^2}{2DC_o}+\frac{h^2}{3D}\right)-\frac{2hC_o}{\pi^2}\sum_{n=1}^{n=\infty}\frac{1}{n^2}\exp\left(-\frac{n^2\pi^2Dt}{h^2}\right)$$
$$+\frac{4C_ih}{\pi^2}\sum_{m=0}^{m=\infty}\frac{1}{(2m+1)^2}\exp\left(\frac{-(2m+1)^2\pi^2Dt}{h^2}\right) \quad (23.2)$$

Eqs. (23.1) and (23.2) show that the initial moisture permeation profiles are not linear at onset, because the exponential terms are not yet negligible, and that the profiles are affected by the initial water concentration in the wall material, regardless of the methods of detection. These equations suggest that the data from the early stage of permeation will vary with the initial conditions and time. Such dependency on the initial conditions cannot be corrected using control containers because the driving forces are different for the test and control containers. Therefore, permeability cannot be determined from the early time point by a single weight gain measurement. However, if the permeation process is allowed to proceed for a sufficiently longer time, such that the exponential terms become negligible, MVTR can be reliably determined using the simplified forms of Eqs. (23.1) and (23.2), regardless of the initial water concentration in the wall and the methods of detection. The effects of initial water concentration on the diffusion profiles, and the advantage of using a steady state method for MVTR can be shown by the following analysis.

For the first method in which the amount of water diffused into the sink is measured, if $C_i=0$, and at $t\to\infty$, the exponential terms in Eq. (23.1) become negligible, and the quantity of water diffused through the wall in Eq. (23.1) can be simplified to a linear function as:

$$Q_{t,d}=\frac{DC_o}{h}\left(t-\frac{h^2}{6D}\right) \quad (23.3)$$

with a lag time t_L being

$$t_L=\frac{h^2}{6D} \quad (23.4)$$

If $C_i=C_o$, and at $t\to\infty$, Eq. (23.1) simplifies to

$$Q_{t,d}=\frac{DC_o}{h}\left(t+\frac{h^2}{3D}\right) \quad (23.5)$$

with a burst time t_B, which can also be called a lag time, being

$$t_B=-\frac{h^2}{3D} \quad (23.6)$$

Eqs. (23.3)–(23.6) show that the initial water concentration in the wall can change the initial diffusion profile to appear either as a lag or as a burst, even though the method of determination remains the same. However, the rate of diffusion over steady state, $dQ_{t,d}/dt=DC_o/h$, is not affected by the initial wall conditions.

For the second method, by which the total weight gain of the entire container is measured, when $C_i=0$, and at $t\to\infty$, Eq. (23.2) can be simplified to a linear function including a burst time in:

$$Q_{t,wg}=\frac{DC_o}{h}\left(t+\frac{h^2}{3D}\right) \quad (23.7)$$

When $C_i=C_o$, and at $t\to\infty$, weight gain simplifies to a linear function with a lag time in:

$$Q_{t,wg}=\frac{DC_o}{h}\left(t-\frac{h^2}{6D}\right) \quad (23.8)$$

Eqs. (23.3), (23.5), (23.7), and (23.8) show that the permeation rates, which are the slopes of the linear lines, are the same as Eq. (23.9) for different methods of detection and for different initial conditions of the

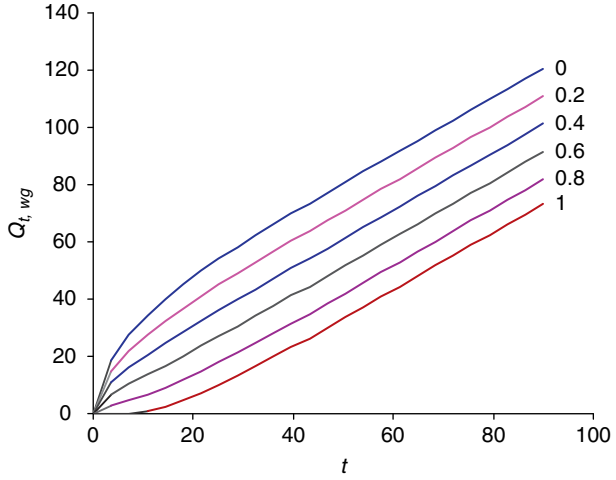


FIGURE 23.1 Simulated weight gain curves of containers filled with desiccants. Numbers at the end of the curves are the values of C_i/C_o. The x and y scales are arbitrary.

wall material, as long as steady state permeation is reached.

$$\frac{dQ}{dt} = \frac{DC_o}{h} \tag{23.9}$$

The definition and the measurement of steady state are discussed in chapter "Fundamental of Diffusion and Dissolution."

The constant rate, by different methods for different initial conditions, show that reliable container permeability can be determined from the steady state permeation data.

The dependency of weight gain profiles on the initial water concentration in the wall, and the advantages of using steady state weight gain data for MVTR, is further depicted in Fig. 23.1 for Eq. (23.2). The figure simulates the effects of initial conditions, to show that the initial weight gain profiles vary from a burst when the wall is initially completely dry, to a delay when the wall is initially equilibrated with external moisture under the test conditions, per Eqs. (23.7) and (23.8). However, if the permeation process is allowed to proceed for a sufficiently longer time to achieve steady state, rate of weight gain (slope of the linear line) is the same regardless of the initial water concentration in the wall if temperature and the driving force remain unchanged. A similar graph can also be constructed, but not shown, for Eq. (23.1), to show that the initial condition has no effect on the rate of permeation through the wall by net permeation if steady state is achieved.

Experimental data for weight gain for two sizes of HDPE bottles containing anhydrous calcium chloride, and closed with induction seal closures with paper backing in the closures, are shown in Fig. 23.2.

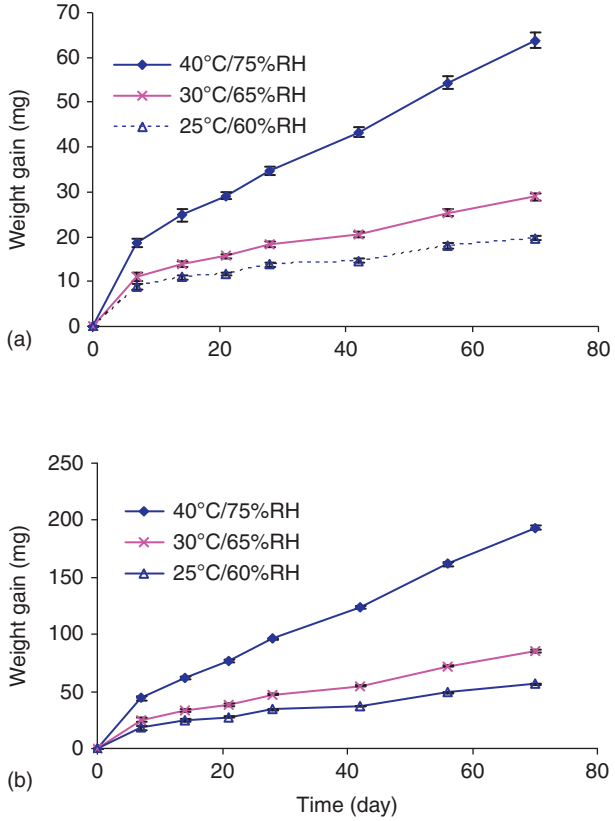


FIGURE 23.2 Weight gain profiles of HDPE bottles containing anhydrous calcium chloride and stored under ICH conditions. Bottle size: (a) 1 oz. (b) 4 oz.

The weight gain profiles agree well with the profiles predicted by theory.

In practice, water concentration in the container wall is not measured. However, it can be assumed that, at the outer surface, water concentration is proportional to the equilibrium external humidity, as measured by relative $\%RH$, by partition at a given temperature. Taking the container surface area into consideration, it can be shown that MVTR during steady state can be described by Eq. (23.10).

$$MVTR = \frac{AdQ}{dt} = \frac{ADK \times \%RH}{h} = \frac{AP \times \%RH}{h} \tag{23.10}$$

where A is the surface area, dQ/dt represents the rate of moisture permeation, D and K are the diffusion coefficient and the partition coefficient of moisture in the wall material, respectively, and P represents the apparent moisture permeability of the wall material.

Other factors to consider for ensuring that the measured container MVTR is reliable include maintaining a "sink" condition, balance sensitivity, and test conditions of humidity and temperature. A perfect "sink" condition is one in which an internal relative humidity of 0%RH can be maintained using a sufficient amount

of effective desiccant in the sealed containers. In real-life tests, maintaining no more than 5%RH of the internal relative humidity during the test may be considered acceptable. Suitable desiccant should be used to achieve a "sink" condition. If anhydrous calcium chloride is used, the desiccant must be activated at a temperature greater than 200°C to remove any potential hydrates on the surface of the desiccant. The current USP Chapter <671> recommends drying of anhydrous calcium chloride at 215°C for 7.25 hours. The moisture sorption capacity of anhydrous calcium chloride in the range below 10%RH at 25°C is less than that of silica gel. Therefore, a large amount of anhydrous calcium chloride must be used to assure the existence of a "sink" condition. Alternatively, some other types of desiccant with a higher moisture sorption capacity than calcium chloride at low humidity, such as silica gel or molecular sieve, may be used in a smaller quantity to achieve a "sink" condition.

For high barrier containers, a sensitive balance should be used to ensure that the variability caused by weighing is negligible compared with the measured MVTR value. The use of aggressive testing conditions in terms of temperature and humidity is helpful to achieve obvious weight changes over the test duration, and thus minimize the effect of weighing error on the MVTR results. The current USP Chapter <671> recommends the use of 40°C/75%RH as the testing conditions for the barrier protection determination of high barrier materials. However, for low barrier materials, such as PVC, the testing conditions should be mild to prevent the saturation of desiccants during testing. ICH conditions must be used if the MVTR data is intended for prediction of product moisture uptake during shelf life under the same conditions. Sample size, test interval, and balance sensitivity should be evaluated together with the estimated container MVTR prior to testing in order to minimize the % variability of the measured MVTR.

MVTRs of some HDPE bottles determined by steady state permeation method are listed in Table 23.3.[15]

TABLE 23.3 Moisture Vapor Transmission Rate (MVTR) and Moisture Permeability of Some HDPE Bottles

	MVTR (mg/day, mean ± SD, $n = 10$)			$P \times 10^6$ (mg · cm/(day · %RH · cm^2))		
Bottle size	25°C/60% RH	30°C/65% RH	40°C/75% RH	25°C	30°C	40°C
40 cc	0.15 ± 0.01	0.26 ± 0.01	0.70 ± 0.02	3.99	6.40	14.67
0.75 oz	0.17 ± 0.01	0.28 ± 0.01	0.71 ± 0.02	3.84	5.88	13.13
1 oz	0.21 ± 0.01	0.34 ± 0.01	0.92 ± 0.01	4.16	6.34	14.71
1.5 oz	0.27 ± 0.01	0.41 ± 0.01	1.02 ± 0.03	4.45	6.30	13.61
3 oz	0.36 ± 0.01	0.53 ± 0.02	1.28 ± 0.02	5.05	6.81	14.36
4 oz	0.59 ± 0.01	0.93 ± 0.03	2.36 ± 0.03	5.19	7.57	16.63
5 oz	0.68 ± 0.02	1.12 ± 0.03	2.82 ± 0.05	5.11	7.76	16.98
175 cc	0.56 ± 0.02	0.94 ± 0.02	2.39 ± 0.04	4.29	6.65	14.70
			Average	4.51	6.71	14.85
			SD	0.53	0.65	1.34

23.2.3 Gas absorbers

23.2.3.1 Desiccants and fillers

Desiccants are commonly used to keep products dry and stable. Dry desiccants can absorb moisture from air either by physical adsorption or by chemical reaction, and thus reduce the humidity in the headspace of sealed containers. Moisture sorption by silica gel is an example of physical adsorption, and sorption by calcium oxide is an example of a chemical reaction. Different types of desiccants and their moisture sorption capacities can be found in literature.[16] The most commonly used desiccants for solid pharmaceutical products are silica gel, clay, and molecular sieves.

Silica gel is an amorphous form of silica (SiO_2 xH_2O) and is highly porous. Silica gel particles are composed of an interconnecting network of microscopic pores (capillaries), and therefore have a very large surface area. The mechanisms of moisture adsorption by silica gel include surface adsorption and capillary condensation in the porous network. Silica gel works well at ambient temperature, but may have decreased adsorption rate and equilibrium moisture content at higher temperatures. The moisture in silica gel can be removed by drying at a temperature greater than 110°C. The current USP Chapter <671> recommends pre-drying silica gel desiccant at approximately 150°C to ensure the complete removal of adsorbed water.

Clay is a low cost and efficient desiccant at low temperature. The primary chemical composition of clay includes silica and aluminum oxide, magnesium oxide, calcium oxide, and ferric oxide. Moisture adsorption capacity may be different for different grades of clay desiccant. Therefore, their moisture adsorption isotherms must be verified during the desiccant selection process. Clay works well at ambient temperature, but at temperatures >50°C, it will likely lose moisture rather than absorb moisture. Similar to silica gel, clay can be dried at a temperature greater than 110°C.

A molecular sieve is a highly porous crystalline material with precise mono-dispersed pores into which certain sizes of molecules can fit, and thus, it can be used to separate one type of molecule from others.

These pores are the channels in the crystalline structure of a material such as zeolite. The pore sizes of molecular sieves differ from silica gel, in that the pore size of a molecular sieve is small and precise, while that of silica gel is much larger with a broad distribution. Molecular sieves are available with different effective pore sizes, such as 3, 4, 5, and 10 Å. A molecular sieve with a 3 Å effective pore size can selectively adsorb water molecules, since the diameter of water molecule is approximately 3 Å, while a molecular sieve with 4 Å pore size can also adsorb nitrogen and oxygen, in addition to water. Molecular sieve desiccants have a very strong affinity and a high adsorptive capacity for water in an environment of low water concentration. At 25°C/10%RH, molecular sieves can adsorb water to approximately 14% of their own weight. This property makes it possible to create an extremely low humidity environment with a small amount of material. However, at any humidity greater than 50%RH at 25°C, the adsorption capacity of a molecular sieve is less than that of silica gel. For example, at 70%RH/25°C, a molecular sieve can absorb approximately 22% of moisture, while silica gel can absorb approximately 32% of moisture at the same condition.[16]

Moisture sorption isotherms of different types of desiccants can be found in literature.[16] The equilibrium adsorptive capacities of silica gel and clay for moisture at different temperatures and humidities are shown in Fig. 23.3. At humidities lower than 35%RH, moisture sorption capacities are similar for the two desiccants. At higher humidities, silica gel has a higher moisture capacity than clay. It is important to know that different grades of the same type of desiccant may have significantly different moisture sorption capacities. This holds true for clay and zeolite desiccants. Therefore, moisture sorption isotherms of desiccants must be verified for the selected desiccants. In addition, the moisture sorption capacity of desiccants is a function of temperature. This temperature dependency should be characterized during the selection process, particularly when the isotherm is intended for modeling packaging design for moisture protection at different temperatures.

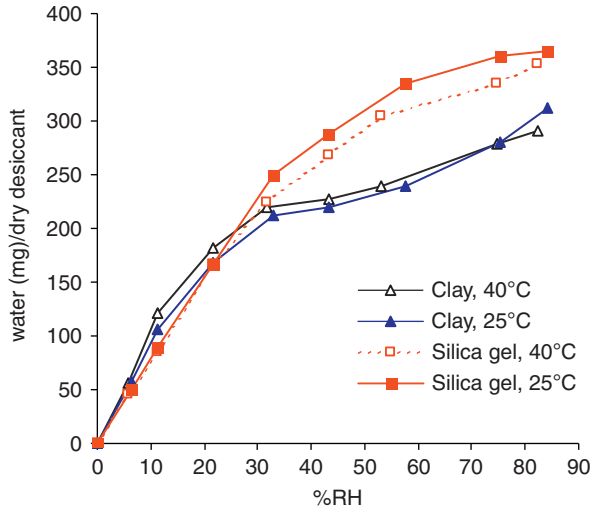


FIGURE 23.3 Moisture sorption isotherms of activated silica gel and bentonite clay.

The common process for using desiccants is to place a predetermined quantity of activated desiccant into containers with products and seal the containers. The quantity of desiccant to be used is important because an insufficient quantity will not provide the required protection, while excessive use of desiccant may lead to over-drying and an unnecessary increase in product cost. In many cases, overuse of desiccant is not a problem for product quality. However, over-drying of some hydrates may lead to the formation of unstable amorphous materials, and thus become detrimental to product quality. It is therefore important to understand the product characteristics, the degradation mechanism of the products, and the desired range of humidity for products before an appropriate quantity of desiccant can be determined. Once the desired range of humidity is determined, a suitable quantity of desiccant to use can be calculated based on the moisture sorption properties of both the desiccant and the drug product using a modeling method, as discussed later in this chapter.

The environmental conditions and packaging process for placing desiccants into product containers must be well-controlled in order to maintain the effectiveness of the desiccant because it may adsorb water vapor quickly when exposed to room air, thus reducing its protection capacity. For example, activated clay, silica gel, and molecular sieves can all absorb approximately 10% of water when exposed to 25°C/75%RH for 1 hour,[16] leading to a significant loss of protection effect for products. In fact, such an exposure can result in a complete loss of the protection effect of silica gel and clay if the desiccants are intended to maintain humidity lower than 20%RH. For container MVTR measurement, adsorption of water by exposure of desiccant to room air prior to filling into containers can lead to underestimation of container MVTR because the internal humidity may exceed the requirement for a "sink" condition. Environmental humidity and exposure time for placing desiccant into containers must clearly be minimized, in order to prevent loss of desiccant effect. Similarly, when active packaging is used where desiccants or oxygen scavenger are pre-sealed in the containers, such as in the bottle caps, the manufacturing, shipping, and handling processes for the containers must also be well-controlled to prevent the loss of the effectiveness of the active packaging.

Care should also be taken when desiccants are used for gelatin capsules to prevent brittleness from loss of moisture. Gelatin capsule shells need approximately 12–16% water to maintain their physical strength, which corresponds with a relative humidity of approximately 40–65%RH at 25°C. Humidity lower than 30%RH will likely cause the capsules to become brittle.[17,18]

Fillers such as cotton and rayon are also commonly copackaged with solid pharmaceutical dosage forms to restrain product movement during shipping. Cotton or rayon can sometimes cause instability problems for drug products. The reasons for this are that these fillers may contain residual oxidative agents that were used for their processing, and that fillers can unintentionally act as either a source of water or as a desiccant. The residual oxidative agents can cause degradation of some active pharmaceutical ingredients (APIs) or result in cross-linking of gelatin capsules or gelatin-coated tablets, if residual formaldehyde exists in the filler, leading to dissolution problems for the drug products. The moisture sorption capacity of cotton is another aspect to consider. The equilibrium moisture content of cotton is showed in Fig. 23.4. It can absorb more than 6% of water at humidity higher than 60%RH at 25°C. If dry cotton is used, it can act as a type of desiccant, and absorb water from the copackaged products. If equilibrated at high humidity during storage prior to use, cotton can be a source of water for the copackaged dry products if the products are hygroscopic. Although the amount of cotton used for pharmaceutical products is small, the effects of cotton on moisture transfer may still be significant in some cases. Quantitative evaluation of the effect of cotton on product moisture content can be carried out using a modeling method, as discussed later in this chapter.

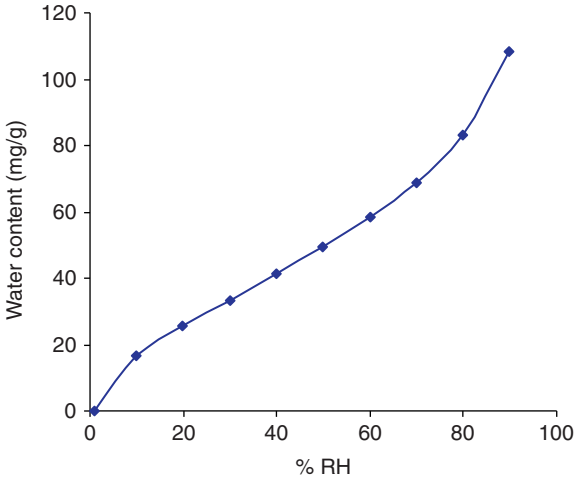


FIGURE 23.4 Moisture sorption isotherm of cotton at 25°C.

23.2.3.2 *Oxygen scavenger*

When products are sensitive to oxygen in the air, one could chose to either:

1. Formulate with antioxidants to stabilize the products;
2. Package the products in containers impermeable to oxygen, such as glass or aluminum containers, combined with purging the container headspace using an inert gas, such as nitrogen or argon;
3. Use oxygen scavengers in well-sealed containers;
4. Combine the techniques of stabilization by formulation and packing protection.

Stabilization by formulation is a complex subject, beyond the scope of this chapter, and will not be discussed here. Once the product is formulated, appropriate packaging can provide further protection of the product to maximize its shelf life. Normal air contains approximately 21% oxygen, and 78% nitrogen. Purging containers with inert gas can reduce the oxygen level to a range of approximately 0.5–5%. The use of a sufficient amount of oxygen absorbers can reduce oxygen content in the headspace of a sealed container to less than 0.1%. Combination of purging to reduce the initial level of oxygen, and the subsequent use of oxygen scavenger in a well-sealed container with low oxygen permeability, is probably the best method for packaging oxygen-sensitive products.

Oxygen scavengers, or oxygen absorbers, are materials that can react with oxygen from the air, eliminating the effects of oxygen when copackaged with products in closed containers. Oxygen scavengers are commonly used for food products, but are less common for pharmaceutical products. In the food industry, different types of materials used as oxygen scavengers include:

1. Inorganic oxygen absorbers, such as powdered metals or their intermediate oxides;
2. Organic antioxidants, such as vitamin C or salicylate; and
3. Polymer-based oxygen absorbers.

The mechanism of product protection by oxygen scavengers is the rapid oxidation of the scavengers with oxygen from the headspace of closed containers, and with the diffused oxygen during storage. According to Grattan and Gilberg,[19] the main reactions for the oxidation of powder iron with oxygen in the presence of moisture and sodium chloride, as in the commercially available oxygen scavenger Ageless, can be described by Eqs. (23.11)–(23.14).

$$Fe \rightarrow Fe^{2+} + 2e^{-} \tag{23.11}$$

$$O_2 + 2H_2O + 4e^{-} \rightarrow 4OH^{-} \tag{23.12}$$

$$Fe^{2+} + 2H_2O \rightarrow FeOOH + 3H^+ + e^- \text{ (rapid)} \quad (23.13)$$

$$3Fe^{2+} + 4H_2O \rightarrow Fe_3O_4 + 8H^+ + e^- \text{ (slow)} \quad (23.14)$$

The rate of oxidation of powder iron in air, and thus oxygen reduction in sealed containers, is rapid. It has been shown that the oxygen level was reduced from approximately 21% to less than 0.1% in approximately 20 hours by using one packet of Ageless Z-1000 in a 2L sealed container.[20] Due to its effectiveness and low cost, metallic iron is widely used to preserve food and museum items.

There are several disadvantages of using metal iron as an oxygen scavenger for pharmaceutical products.

1. The reaction of metal iron with oxygen requires a water activity as high as 65–85%RH in the system.[19] In fact, water was included in the form of saturated sodium chloride solution absorbed in molecular sieves as part of the oxygen scavenger, to maintain a critical relative humidity for reaction.[19] A high humidity environment may be detrimental for most solid pharmaceutical products.
2. The reaction is exothermic. The heat released by the reaction may create thermal stress for the copackaged products.
3. The reaction creates a vacuum due to the elimination of oxygen in the container, which may cause flexible containers to collapse. Therefore rigid containers are needed in order to maintain the geometry of product packages using this type of oxygen scavenger.

Unlike desiccants, oxygen scavengers are not yet well studied for pharmaceutical products. Therefore, much study, including the characterization of material properties and the oxidation-reduction mechanisms of oxygen scavengers and drug products, needs to be done to evaluate oxygen scavengers for protection of pharmaceutical products. Successful oxygen scavenger candidates for products must be effective, safe to use, clean (so as not to contaminate products), convenient to use, and low cost.

23.2.4 Drug products

While there are different types of containers available, selection of containers for product protection clearly depends on the specific protection requirements of any given product. In general, the protection requirements are identified during the early development stage with accelerated degradation studies or forced degradation studies using either the APIs or drug products. Details on stability studies were discussed in chapters "Drug Stability and Degradation Studies" and "Excipient Compatibility and Functionality." Once the drug sensitivity and protection requirements are identified, container design and selection may be made by matching the properties of containers with that of the drug products. If accelerated studies show that the product is not sensitive to moisture and/or other gases, the protection requirements for the product will mainly include the prevention of contamination and physical stress. If a product is sensitive to light, containers with low light transmittance should be the first choice, otherwise, secondary packaging will be needed if transparent primary packaging is used. If a product is sensitive to moisture, containers with low MVTR values are preferred, and desiccant may be used for additional protection.

Different pharmaceutical excipients have different moisture sorption capacities. The same is true for finished pharmaceutical products, since each product contains its own set of excipients. Product moisture sorption capacity is one of the critical factors that determine the rate and extent of moisture ingression into containers during storage, in addition to the container MVTR and storage conditions. The higher the capacity, the faster the ingression rate will be, assuming other factors are the same. If desiccants are used for products, it is important to understand the solid state chemistry of the APIs in the products. Some APIs may exist in different hydrates and/or crystal forms under different humidity. Over-drying caused by use of excessive desiccants may lead to lose of hydrates, and formation of unstable lower hydrates or amorphous materials. Therefore, drug product properties, including moisture sorption isotherm and the solid state response to environmental conditions, should be studied. This is particularly important if desiccants are used for quantitative design of packaging to protect moisture-sensitive products. The application of material moisture sorption isotherms and container moisture permeability in quantitative packaging selection for moisture-sensitive products are discussed in the next section.

23.3 LINKING PACKAGING PROPERTY WITH DRUG PROPERTY

MVTR is traditionally used to classify containers, and the use of desiccant for products is determined by trial and error. According to USP Chapter <671> prior to USP 37th edition, there was only one method using MVTR for classification, containers are *tight*

containers if not more than one of the 10 test containers exceeds 100 mg/day per L, and none exceed 200 mg/day per L. Containers are *well-closed containers* if not more than one of the 10 test containers exceeds 2000 mg/day per L, and none exceed 3000 mg/day per L. This type of classification is very loose, and does not take into account the properties of products in the containers. As a result, this classification is not meaningful in quantitative packaging design for the support of new product development. This problem has been acknowledged by the pharmaceutical industry and by USP. Therefore, new methods are now included in the recent USP38 Chapter <671> to provide reliable and precise MVTR of containers for differentiation of container performance in the section regarding barrier protection determination for packaging systems for SOD forms.

In product research and development (R&D), we must know if any container can provide adequate protection of a new product, prior to the use of the container for official stability studies. Therefore, vigorous evaluation is required for container selection. The conventional method of trial-and-error for packaging and desiccant selection is of low efficiency, and lacks a guiding power. A desirable approach should be science-based, and should allow the building of QbD. Two new methods have been developed in recent years to evaluate container moisture barrier properties for products. These methods are the use of MVTR/unit dose product,[13,21] and the modeling approach for prediction of moisture uptake by packaged products during storage.[22] The key element of these two new methods is the linking of packaging property with drug product property, and they are discussed in the next section.

23.3.1 The use of moisture vapor transmission rate per unit product for container comparison

For a given product in different containers, the moisture protection efficiency of the containers for the product can be evaluated using a criterion of "MVTR per unit quantity of product."[13,21] For solid products, the unit can be a tablet, a capsule, or the mass of a product. A smaller value of MVTR/unit product indicates that the container provides better protection for the product from moisture than a container with a larger value. This criterion can be used to rank containers/designs, and to guide packaging selection for new product R&D, or it can be used to evaluate the container moisture barrier equivalence for justification of postapproval container changes for a regulatory purpose.

23.3.2 Modeling of moisture uptake by packaged products

Moisture content of products in permeable containers stored at ICH stability conditions can be predicted by a modeling method[22] if it is assumed that:

1. Moisture content of a product is a function of the equilibrium humidity;
2. Moisture permeation through the container is the rate-limiting step;
3. A quasi-steady state of permeation exists; and
4. Moisture permeability of the container is constant at a given temperature.

At quasi-steady state, the rate of moisture permeation through a sealed container at constant temperature and external humidity can be expressed using Fick's first law as:

$$\frac{dQ}{dt} = k(RH_{\text{out}} - RH) \tag{23.15}$$

where

dQ/dt is the rate of moisture permeation through the container

k represents the apparent moisture permeability of the container

RH_{out} and RH represent the % relative humidity outside and inside the container, respectively.

If the water content of any unit component in the container is a function of the equilibrium humidity, the total quantity of water, Q, in a sealed package containing n components can be written as:

$$Q = \sum_{i=1}^{i=n} q_i f_i(RH) \tag{23.16}$$

where q_i and $f_i(RH)$ are the quantity and the moisture sorption isotherm of the ith component in the container, respectively.

Differentiation of Eq. (23.16) with respect to relative humidity leads to

$$\frac{dQ}{d_{RH}} = \frac{d}{d_{RH}}\left[\sum_{i=1}^{i=n} q_i f_i(RH)\right] \tag{23.17}$$

By mass balance, the quantity of moisture permeated through the container must be equal to the moisture content change in the container, as described by Eq. (23.18).

$$k(RH_{\text{out}} - RH)d_t = \frac{d}{d_{RH}}\left[\sum_{i=1}^{i=n} q_i f_i(RH)\right] d_{RH} \tag{23.18}$$

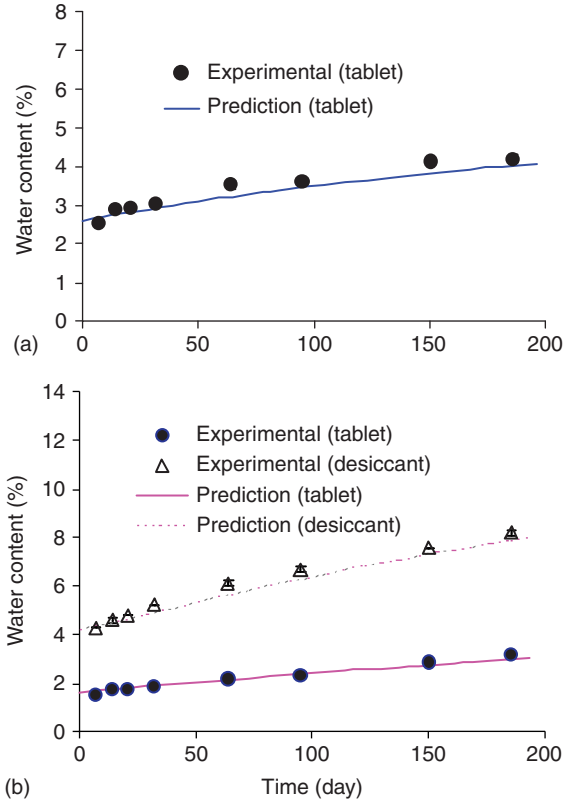


FIGURE 23.5 Experimental and predicted water content of 20 model tablets in sealed 3 oz HDPE bottles and stored at 40°C/75% RH. (a): no desiccant. (b): with 1 g bentonite desiccant.

Rearranging and integration leads to the *t-RH* profile inside the container as described by Eq. (23.19).

$$t = \int_{RH_{t=0}}^{RH_t} \frac{\frac{d}{d_{RH}}\left[\sum_{i=1}^{i=n} q_i f_i(RH)\right]}{k(RH_{out} - RH)} d_{RH} \qquad (23.19)$$

The water content of each individual component as a function of time can then be estimated by substituting the *RH* at time *t* into the corresponding moisture isotherm of the component, as given by the *t-RH* profile. Detailed procedures for using this modeling method can be found in literature.[22] Predicted moisture content for a model tablet product in HDPE bottles stored under accelerated and long-term ICH conditions agreed very well with experimental data (Figs. 23.5 and 23.6), demonstrating that this modeling method is highly reliable. The reliability makes it possible to practice a QbD approach in packaging selection for products. Using this modeling method, the effects of any packaging designs and storage conditions on product moisture content can be readily simulated to select the most suitable packaging for moisture protection. The selected packaging can then be confirmed

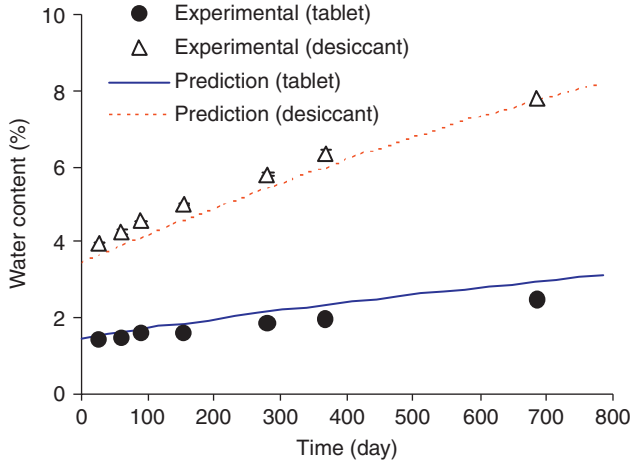


FIGURE 23.6 Experimental and predicted water content of 20 tablets in a sealed 3 oz HDPE bottle with 1 g bentonite desiccant and stored at 25°C/60%RH.

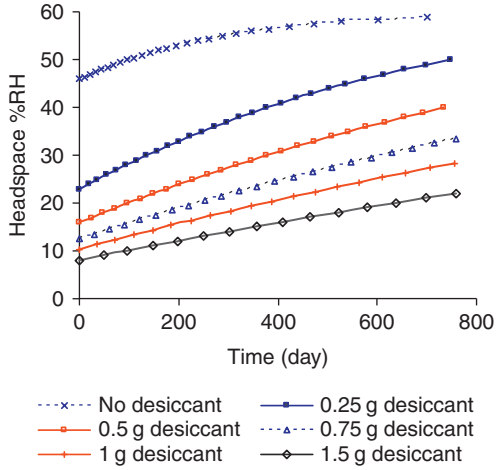


FIGURE 23.7 Simulation of headspace humidity for the same amount of a model product in a 1 oz HDPE bottle with different quantity of silica gel and stored at 25°C/60%RH.

by real-life stability study. Suppose it is known that the headspace humidity of a product must not be greater than 30%RH at room temperature over its shelf life, and that desiccant is to be used in HDPE bottles for the product. However, it is not known what quantity of desiccant should be used and what the effect of moisture permeation during the shelf life will be. With the moisture sorption property of the product, and the container permeability at hand, simulation can be conducted for different amounts of desiccant. The simulation results are shown in Fig. 23.7. Based on the simulation, it can be determined that at least 1 g of the desiccant must be used, in order to maintain the headspace relative humidity and not to exceed 30%RH over 2 years when stored at 25°C/60%RH. In another case, 7 and 90 tablets were packaged

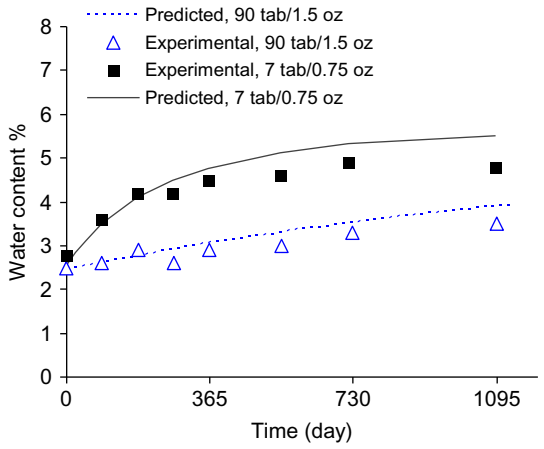


FIGURE 23.8 Predicted and experimental moisture content of tablets in different designs using HDPE bottles and stored at 25°C/60%RH.

without desiccant in 0.75 and 1.5 oz HDPE bottles, respectively. Simulation indicated that the moisture content of the 7 tablets in the 0.75 oz bottle would approach an equilibrium value in approximately 1 year when stored at 25C/60%RH, while that of the 90 tablets in the 1.5 oz bottle would not reach the same level of moisture content over 3 years. The simulated results were confirmed by formal stability study conducted under ICH conditions (Fig. 23.8). These examples show that the modeling method can be used as a scientific tool to evaluate the moisture protection capability of any proposed packaging design, prior to conducting stability studies for new product development, or to evaluate the container equivalence for postapproval container changes. Formal stability studies can then be conducted for confirmation purposes. Modeling is science-based, and can be used as an effective tool for QbD in packaging evaluation for product protection against moisture. Successful application of this method will reduce product development's time and cost.

23.4 POSTAPPROVAL PACKAGING CHANGES

Postapproval changes to packaging materials (PACPAC), including a change to a container closure system, to a component of the container closure system, to a material of construction for a component, or to a process involving one of these, must be reported to the regulatory agency.[2,23] Packaging changes are classified as either major, moderate, or minor, depending on the nature of changes. The categories of changes indicate the regulatory requirements, reporting categories, and thus the levels of burden on the FDA and on the pharmaceutical industry. Major changes require a prior approval supplement. Moderate changes need a changes being effective supplement, while minor changes can be reported to the FDA by annual report. Categorization of changes depends on the nature of the changes, and the intended use for products, such as for injectable, ophthalmic, liquid oral, solid oral, topical, or transdermal dosage forms. It should be understood that the degree of concern for product-packaging interaction, and thus categorization of packaging changes, is different for different types of dosage forms. For example, changes in the size or shape of a container for a sterile drug product may be considered as moderate changes, while the same changes might be treated as minor changes for SOD forms. Therefore, a category of PACPAC for one type of product cannot be assumed to be the same category for another type of product.

For SOD forms, there are several questions to answer before proposing postapproval packaging changes. These include, but are not limited to, the following;

1. Is the new packaging previously approved by the FDA for the same type of dosage form?
2. What type of protection does the product need?
3. Does the new packaging provide equivalent or better protection?
4. What type of packaging material information and what type of product stability testing in the new packaging are needed for submission?

Current FDA guidance to industry, changes to an approved new drug application or abbreviated new drug application,[23] provide some examples for each category of packaging changes and a general statement on the potential impact the changes may have on the identity, strength, quality, purity, or potency of a drug product as they may relate to the safety or effectiveness of the product. However, the guidance does not list specific requirement on information needed to report to the FDA for the changes. A regulatory guidance listing specific instructions on what and how to report for postapproval changes for packaging will be helpful to reduce the regulatory burden for the FDA and the pharmaceutical industry.

The reporting category of packaging changes may be reduced if the new packaging has been approved by FDA for the same type of dosage form, and if it is demonstrated that the new packaging can provide equivalent or better protection than the existing packaging. If it can be demonstrated that a product is not sensitive to environmental conditions, such as moisture, oxygen, or light, then the need to evaluate the protection equivalency is not applicable for packaging changes. Protection equivalency can be demonstrated either by formal stability studies or by comparative tests. A formal stability study is not preferred for

material characterization because it is time consuming and expensive. Thus, comparative tests are preferred to increase efficiency and reduce cost. For SOD forms, container MVTR is commonly used as one of the criteria to demonstrate the moisture protection equivalency of containers. Two aspects of MVTR that should be considered are the reliability of MVTR values and the correct use of those values. As discussed previously in this chapter, the current USP chapter <671> provides different methods for measuring the container MVTR for either container classification or for barrier protection determination purpose. Results of the classification method may be highly variable, and less reliable than that of the method for barrier protection determination. Thus, the author prefers the latter method to obtain container MVTR for container comparison.

Getting a reliable MVTR value does not yet provide the basis for evaluation of container equivalency for product protection. The correct use of the MVTR is equally important for drawing meaningful conclusions. For classification purposes, MVTR is reported in a unit of mg/L per day as per USP. This unit has limited utility, because it does not take into account the quantity of products in the containers. Alternatively, meaningful results can be obtained by using a new unit of "MVTR per unit product"[13,21] proposed by the PQRI container closure working group. According to the studies conducted by the PQRI workgroup, moisture activity, and hence product moisture content in sealed containers, is related to the value of "MVTR per unit product." The lower the "MVTR per tablet" value, the lower the product moisture content, and it does not matter if the container is an HDPE bottle or an Aclar unit blister. Several examples have been reported in previous publications by PQRI. More examples for using "MVTR per unit product" to rank container moisture protection property, and to link it to product moisture content by stability studies are listed in Table 23.4. Tablets of a same lot were packaged in unit blisters, and in two sizes of HDPE bottles, for formal stability study at 25°C/60%RH. The ranking of the MVTR/unit for the three types of packaging are in the order of blister > bottle 1 > bottle 2. At the 1-, 2-, and 3-year time points, significant differences in tablet moisture content were observed in the order of - blister > bottle 1 > bottle 2. These data clearly demonstrate that "MVTR per unit product" is a valuable criterion for evaluating moisture protection equivalency of containers.

To summarize, the unit of "MVTR per unit product" links container property with product quality. The unit provides a means of evaluating different types of packaging materials, such as bottles and blisters, as well as for comparison of different quantities of products in varying or the same sizes of containers. Protection capability is the same or better, as long as "MVTR per unit product" is the same or smaller if moisture protection is the main concern, regardless of the types of packaging materials and configurations. MVTR per unit product is therefore a science-based criterion for evaluating the moisture protection equivalency, and for justifying postapproval packaging changes.

TABLE 23.4 Ranking of Moisture Vapor Transmission Rate per Tablet and the Resulting LOD of a Model Product

Packaging	MVTR/tablet (mg/tab per day at 25°C/60%RH)	Product lot	%LOD at 25°C/60%RH				
			Initial	6M	12M	24M	36M
Unit blister	0.078	1	2.7	3.4	3.8	4.5	4.6
		2	2.8	3.5	3.8	4.5	4.6
		3	2.7	3.4	3.8	4.6	4.6
HDPE bottle 1	0.024	1	2.8	3.2	3.5	3.9	4.1
		2	2.7	3.2	3.3	3.9	4.0
		3	2.7	3.3	3.4	3.9	4.1
HDPE bottle 2	0.004	1	2.6	2.7	2.6	2.7	2.7
		2	2.6	2.8	2.6	2.8	2.7
		3	2.8	2.7	2.7	2.9	2.9

References

1. ICH Harmonied Tripartite Guideline, Q1A(R2). Stability testing of new drug substances and products; November 2003. Q1B photostability testing of new drug substances and products, November 1996; Q1C Stability testing for new dosage forms; May 1997.
2. FDA Guidance for Industry. Container closure systems for packaging human drugs and biologics; 1999.
3. FDA Guidance for Industry. Changes to an approved NDA and ANDA; November 1999.
4. USP 37, <671>.
5. Hanlon JF. *Handbook of package engineering*. 2nd Ed Lancaster, PA: Technomic Publishing Company, Inc.; 1992.
6. Weeren RV, Gibboni DJ. Barrier Packaging as an Integral Part of Drug Delivery. *Drug Del Technol* 2002;**2**:48–53.
7. Armstrong RB. Effects of polymer structure on gas barrier of ethylene vinyl alcohol (EVOH) and considerations for package development. TAPPI PLACE conference, Boston, MA, September 9–12, 2002, p. 243–69.
8. Kamal MR, Jinnah IA. Permeability of oxygen and water vapor though polyethylene/polyamide films. *Polym Eng Sci* 1984;**24**:1337–47.
9. Sezi R, Springer J. Gaspermeabilität von Polypropylen. *Colloid Polym Sci* 1981;**259**:1170–6.
10. Shams Rustom. 2011. <http://www.interphex.com/RNA/RNA_Interphex_V2/documents/2011/handouts/Rustom_Shams.pdf>.

11. Unisource Worldwide, Inc., Specification. <http://www.unisourcelink.com>.
12. Taborsky CJ, Grady LT. Water vapor permeation of prescription vials. *Pharmacopeial Forum* 2001;**27**:2417–27.
13. PQRI Container-Closure Working Group. Basis for using moisture vapor transmission rate per unit product in the evaluation of moisture-barrier equivalence of primary packages for solid oral dosage forms. *Pharmacopeial Forum* 2005;**31**:262–9.
14. Crank J. *The mathematics of diffusion*. Oxford: Clarendon Press; 1986. p. 49–51.
15. Chen Y, Li Y. Determination of water vapor transmission rate (WVTR) of HDPE bottles for pharmaceutical products. *Int J Pharm* 2008;**358**:13–143.
16. Dobson RL. Protection of pharmaceutical and diagnostic products through desiccant technology. *J Packaging Technol* 1987;**1**:127–31.
17. Kontny MJ, Mulski CA. Gelatin capsule brittleness as a function of relative humidity at room temperature. *Int J Pharm* 1989;**54**:79–85.
18. Chang RK, Raghavan KS, Hussain MA. A study on gelatin capsule brittleness: moisture transfer between the capsule shell and its content. *J Pharm Sci* 1998;**87**:556–8.
19. Grattan DW, Gilberg M. Ageless oxygen absorber: chemical and physical properties. *Stud Conserv* 1994;**39**:210–14.
20. Lambert FL, Daniel V, Preusser FD. The rate of absorption of oxygen by Ageless™: the utility of an oxygen scavenger in sealed cases. *Stud Conserv* 1992;**37**:267–74.
21. Bergum J, Chen Y, Hunt D, Lockhart H, Malinowski D, Moreton C, PQRI Container–Closure Working Group, et al. Development and application of MVTR/unit data in regulatory submissions. *Pharmacopeial Forum* (Online) 2013;**39**(6). <www.usppf.com/pf/pub/index.html>.
22. Chen Y, Li Y. A new model for predicting moisture uptake by packaged solid pharmaceuticals. *Int J Pharm* 2003;**255**:217–25.
23. FDA Guidance for Industry. Changes to An Approved NDA or ANDA; 1999.

This page intentionally left blank

CHAPTER

24

Clinical Supplies Manufacture: Strategy, GMP Considerations, and Cleaning Validation

B.W. Pack[1], *S. Stithit*[2] *and W. Chen*[3]

[1]Eli Lilly and Company, Indianapolis, IN, United States [2]Century Pharmaceuticals, Indianapolis, IN, United States
[3]Gilead Sciences, Foster City, CA, United States

24.1 INTRODUCTION

The process of new drug discovery research, development, and commercialization is long, complex, and costly. It can take more than 10 years and cost $1 billion or more on average to bring a new chemical entity (NCE) to the market. The new drug development process can be characterized into different phases including Phases I, II, III, and IV. The objective for each phase is described in Table 24.1. From Phase I to Phase III, clinical investigation is one of the most important activities to prove the safety and efficacy of a new drug. Successful supply of clinical materials is key to assure that clinical studies are conducted as planned at the correct time with the correct dose and correct amount of materials. Manufacturing batch sizes of clinical supplies change with demand, availability of drug substances, and study phases. A strategy clearly must be developed to effectively manage the complex clinical supplies manufacturing process. This includes planning, collaborative communication among the cross-functional team, lean manufacturing, effective training programs, and application of new technologies to facilitate the manufacture of clinical supplies.

Clinical trials are designed to obtain safety and efficacy data for new drug candidates. A primary consideration in these studies is the safety of the clinical trial subjects. In addition, careful consideration must be given to the quality of the clinical supplies produced to ensure that there is no unintended consequence on the clinical trial results. Therefore, a comprehensive cGMP controls must be in place for the manufacture of clinical supplies to ensure the delivery of quality products. A cleaning verification and/or a cleaning validation program is a key component of the cGMP controls and quality system and an essential part of any pharmaceutical manufacturing facility. It must be demonstrated either through cleaning-validation effort that the cleaning process is under control, or through cleaning verification after each manufacture so that there is no cross-contamination from one lot to the next and the safety of product users is assured.

24.2 STRATEGY OF CLINICAL SUPPLIES MANUFACTURE

In the past few years, there have been major challenges facing the pharmaceutical industry regarding growing regulatory burdens and global clinical trials. More rules and regulations on good clinical practice (GCP) have been implemented or recommended because of ongoing concerns regarding unreported adverse events, inadequate protection of the clinical trial subjects (especially minors and incapacitated adults), inadequate oversight of clinical trials by Institutional Review Boards (IRBs),[1] unethical use of placebo-controlled trials, and insufficient informed consent practices. At the same time, the costs of drug development and discovery have increased substantially. Bringing a new drug to market can cost anywhere between $900 million and $1.7 billion, depending on development and marketing costs. Approximately only one in 10,000 compounds synthesized will pass regulatory approval and be commercialized, with only a few becoming potential

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00024-8

TABLE 24.1 Clinical Trials in New Drug Development

Stages	Time	Subject/study	Study objectives
Clinical Trials in Phase I	1–2 years	20–80 healthy volunteers Open study	Safety; Dose and schedule; Pharmacokinetics; Pharmacological actions
Clinical Trials in Phase II	2–3 years	100–200 patients Controlled studies-multicenter	Safety and efficacy; Dose, route and schedule; Toxicity profile; Clinical endpoints; Dose–response curve; Short-term adverse effects; Pharmacokinetics; Pharmacological actions
Clinical Trials in Phase III	2–4 years	>1000 patients Controlled studies-multicenter	Safety and efficacy in a large population; Patient population; Product claims; Final formulation and product stability; Drug-related adverse effects
Clinical Trials in Phase IV	2–4 years	Controlled studies-multicenter	Additional information about the drug's risks, benefits and optimal use such as other patient populations, other stages of diseases, over a longer period, extend claims of safety and efficacy

blockbuster drugs. In addition, the time required to bring a new drug to the market is becoming longer. The time from drug discovery to FDA approval for a NCE can be as long as 12 years. Pharmaceutical companies are being forced to examine how to streamline their manufacturing practices and products.[2] The faster the new drug product is approved by the FDA, the longer the effective life of market exclusivity for the product before its patent expires.

Pharmaceutical manufacturers often rely on various strategies to keep their new drug pipeline flowing freely. Effective management of clinical supply delivery prevents it from becoming a rate limiting step in product development, reduces the time for a new drug undergoing clinical trials, and ultimately increases the speed of bringing a new drug to market. Due to the complexity and long duration of clinical trial programs, success depends on precise timing, monitoring, managing, and coordination efforts as well as a smooth flow of information and materials. The primary goal is to help connect clinical trial management teams with clinical supply manufacturing. The proper planning process is crucial to ensure that the project timeline is met. The collaborative communication among team members including research and development (R&D) personnel, quality control/quality assurance (QC/QA) personnel, regulatory personnel, medical personnel, clinical investigators, clinical trial material manufacturing personnel, and clinical trial material packaging personnel is essential for the success of the project. An implementation of clinical plan, clinical supplies liaison, a lean manufacturing approach, a cross-functional training program for workers, outsourcing, the use of new technology; including modern computer technology and new methodologies in drug discovery and development and process analytical technology (PAT) are among the various strategies that can be used to improve clinical trial management. These topics will be discussed later in this chapter.

24.3 CLINICAL PLAN

For each new drug candidate targeted for eventual registration, marketing, and distribution, a clinical plan is generally required. The clinical plan provides a detailed listing of clinical trials required for FDA approval together with duration and the desired timelines.[3] In general, project plan rationale, workload, and cost are discussed at management meetings or at project team meetings. Planning must include short-term, intermediate-term, and long-term goals with assumptions for bulk drug substance requirements, dose strengths, and quantity of drug product to manufacture. Comparator drug products also need to be considered. The short-term and intermediate-term planning (1–6 months) must include final plans for manufacturing, including the bulk product request with an approximately 4- to 6-month lead time, and packaging requests, which should be received 2–3 months before the expected shipping date. Long-term strategic plans can be prepared based on equipment/facility capacity, volume prediction, budget or resources, preferred vendor or contractor relationships, and comparator or drug substance sourcing. There are many challenges in planning for distribution of the clinical trial supplies to multiple investigation centers, especially those in different countries. Regulation-specific requirements for each country running the clinical trials must be thoroughly understood and considered during planning.

Long-term strategic planning may be based on workload prioritization, contract arrangements, and inventory management. On the other hand, short-term or immediate planning is more related to completing jobs at hand. Clinical plans and process must be reviewed and updated on a regular basis to keep the liaison and other team members in the same pace with the possible changes for modification of protocols. It is generally a good idea to establish a frame of reference to define how, when, and who will change what—commonly referred to as a "change control" plan. This

change control should be complementary to the planning process. Planning is the process that requires management to ensure that the project is completed within the timeline. All the risks involved in the project should be identified and assessed during the planning process. The probability of the risk occurring, the impact of the risks on the project, and a contingency plan should be incorporated into the planning activities.

A clinical plan template can be developed because clinical studies usually follow the same basic tasks and flow, regardless of the product is being studied. Using the template approach helps accelerate the project timeline and bring in best practices learned from past development projects. Table 24.2 illustrates the tasks commonly required in Phase I clinical trial studies.

TABLE 24.2 Template for Phase I Clinical Trials

Activity name	# of days	Start date	End date	Responsible manager
Clinical protocol development • Study objective and rationale • Study design and outcomes • Subject inclusion and exclusion criteria and enrollment procedures • Study interventions • Clinical and laboratory evaluations • Management of adverse experiences • Criteria for intervention discontinuation • Statistical considerations • Data collection/site monitoring • Adverse experience reporting • Subjects/IRB review/ informed consent				
Investigator and site selection • Investigator's brochure • Sponsor-investigator agreement				
Clinical supply vendor/ contractor selection • Plan and management of vendors • Communication plan • Vendor quality assurance/ cGMP audit • Signed agreement				
Clinical trial materials and distribution plan • Manufacturing • Packaging and labeling • Shipment				
Budget/financial aspect				
Regulatory filing / reporting • Trial master file • Study site notebook				
Clinical study management/ monitoring plan				
Case report form (CRF)/data management • Database design and setup • Data Security and validation				
Quality assurance of clinical trial documents/activities • Site GCP audit				
Risk identification and contingency plan				
Training all personnel involved in the study				
Site management				
Subject recruitment initiation				

24.4 CLINICAL SUPPLIES LIAISON

The clinical supplies liaison (or clinical trials materials manager/team leader, or medical coordinator or clinical logistic coordinator) plays an important role in the drug development-clinical trial process. The titles or names of this position may differ, but the tasks that need to be performed are essentially the same. The liaison must understand and be able to successfully lead the project, effectively communicate and convey the clinical study process and progress to team members, and prioritize task activities and timelines. The liaison should understand team member individuals and their cultures. He/she must be able to organize, manage, negotiate with, educate, and compromise with multidisciplinary team members to develop the best scenario for all parties involved. A successful liaison must have a strong scientific background and strong people management skills (ie, the ability to interact and effectively negotiate with other people). A successful liaison can lead his or her team or customers to achieve a sought-after goal and eventually push the project through the finishing line despite numerous setbacks, delay, adversity, and failures.

24.5 LEAN MANUFACTURING

Lean manufacturing is a business management philosophy that seeks to eliminate or reduce waste and maximize customer value through employee involvement. It evolved mostly from successful Japanese practices in production systems. The lean manufacturing paradigm calls for integration of employee involvement and technological practices. It focuses on the human-centered approach to the design and implementation of advanced manufacturing systems.[4,5] Lean

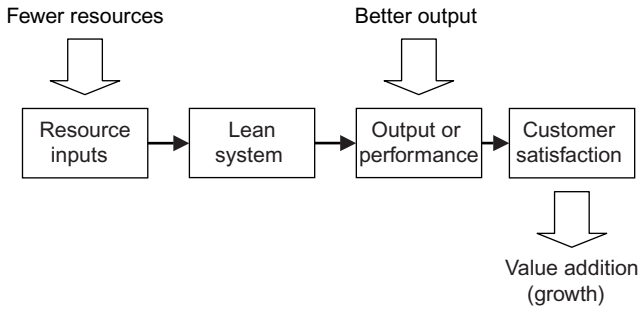


FIGURE 24.1 Lean manufacturing system.

manufacturing is a dynamic system that requires fewer resources, less material, less inventory, less labor, less equipment, less time, and less space. It is intended to eliminate waste while meeting and exceeding customer requirements, and yielding better output (eg, quality, variety, cost, time, and safety) (Fig. 24.1).

There are three main approaches in lean manufacturing that lead to better output. First, lean manufacturing involves a new philosophy of manufacturing, which focuses on better quality, exceeding customer requirements, and continuous improvement. Second, lean manufacturing involves establishment of a new culture and managerial system to incorporate the lean philosophy. The lean team leader and lean steering committee are included in their organizations to manage product development, supply chain relations, production operations, and coordination of the overall enterprise toward lean manufacturing. Third, lean manufacturing uses techniques that address specific problem such as just-in-time (JIT), total productive maintenance (TPM), total quality management (TQM), human resource management (HRM), scientific management, six sigma, quality circles, simultaneous engineering, and inventory systems. In a true lean organization, defects or problems are discovered immediately and solutions are sought at once using whatever techniques are required.

Fig. 24.2 shows a model for implementation of the lean manufacturing in clinical supplies manufacture in a pharmaceutical firm. The first step occurs when management commits to the lean philosophy illustrated by the three goals:

1. Zero waste is to eliminate what does not belong to the value stream;
2. Value Stream Mapping is to identify the scope of activities conforming the value stream; and
3. Pull (Kaizen) is to develop workforce capabilities by continuous, gradual, and persistent improvement.

As a new practice of manufacturing, lean manufacturing demands new capabilities from all employees at all levels: problem-solving focus,

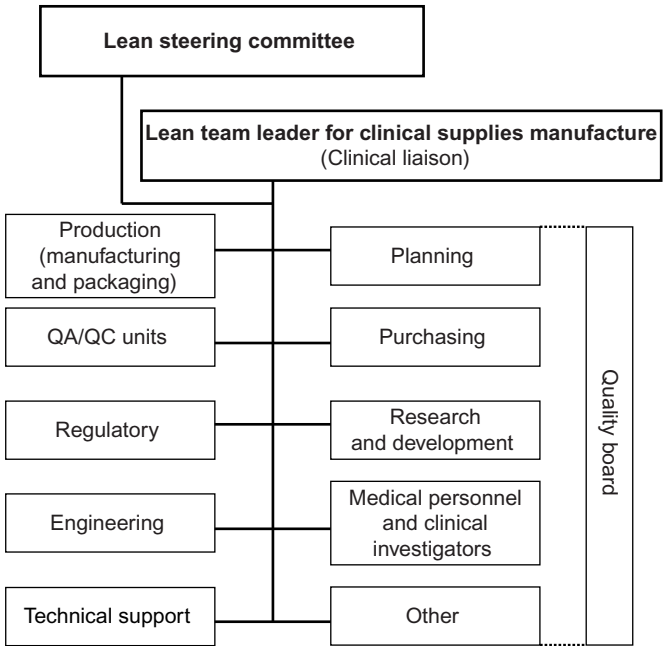


FIGURE 24.2 Implementation of lean manufacturing.

teamwork, and creative thinking. Lean manufacturing is not a fixed set of techniques, as has been discussed previously. As part of the implementation, new techniques unique to each business culture will emerge at each company. However, since businesses share the same interests and join a competitive market, common techniques are likely to appear. The examples of common techniques include:

- JIT: JIT is a philosophy of manufacturing that focuses on supplying the right materials to the right place in the right quantity JIT to achieve perfect work flow while minimizing waste. It depends on the pull process from the consumer (demand) end, not the push from the production (supply) end. Materials are supplied to the production line only when required. JIT can be achieved by using these techniques: lot size reductions, continuous flow production, pull systems, cycle-time reductions, focused factory production systems, agile manufacturing strategies, quick changeover techniques, bottleneck or constraint removal, and reengineered production processes.
- TPM: TPM aims to increase equipment effectiveness by eliminating the losses due to machine downtime using techniques such as predictive or preventive maintenance, maintenance optimization, safety improvement programs, planning and scheduling strategies, and new process equipment. TPM works on continuous flow of production. Any interruption to the flow such as machine breakdown is considered and treated as waste.

- TQM: TQM is designed to improve customer satisfaction and product quality by using several techniques such as quality systems, quality management programs, process capability measurements, six sigma, competitive benchmarking, and formal continuous improvement programs.
- HRM: HRM is designed to effectively use human resources to increase overall performance. It focuses on leadership, empowerment, communication, and training by using several techniques such as self-directed work teams, cross-functional frameworks, and open communications in order to bring out the best possible employee capabilities.

Clinical supplies manufacture provides cGMP compliant products at a capability that is designed to deliver on-time clinical supplies meeting or exceeding customer needs in support of clinical trials. Clinical supplies for the early phase clinical trials (Phases I and II) typically need to be acceptable to good stability, meet basic dosage form requirements (ie, assay, impurities, content uniformity), and specific release requirements (ie, dissolution), and be suitable for small to intermediate scale manufacturing (ie, from hundreds to thousands of units or from several to tens of kilogram). On the other hand, clinical supplies for the later phase clinical trials (Phase III) typically need to have good stability, meet all dosage form requirements (including in-process testing requirements) and specific release requirements (ie, dissolution), and be suitable for larger pilot scale manufacturing (ie, $\sim$50–100 kg) and scalable to production scale ($>$100 kg). The numbers of products and volume to manufacture also change as the study phase advances. In general, the number of products to manufacture and thus process variety for clinical supplies decreases and product volume increases with the advancement of clinical trial phases (see Fig. 24.3).

The JIT technique can be effectively applied to clinical supplies manufacture for early phase clinical trials because of the high variety and low product volume for this phase. When clinical trials advance into Phase III, lean manufacturing techniques, including JIT, total preventive maintenance, TQM, and HRM, can be applied to the clinical supplies manufacture because of low variety and high volume of the batches.

The lean manufacturing philosophy and methodology can be applied throughout the spectrum of development and manufacturing, from preclinical to clinical, from the overall strategic level of the development operation to the specific and detailed level of the manufacturing processes. The lean manufacturing provides a structured, disciplined, and logical progression for organizations to achieve quality improvement and breakthrough.

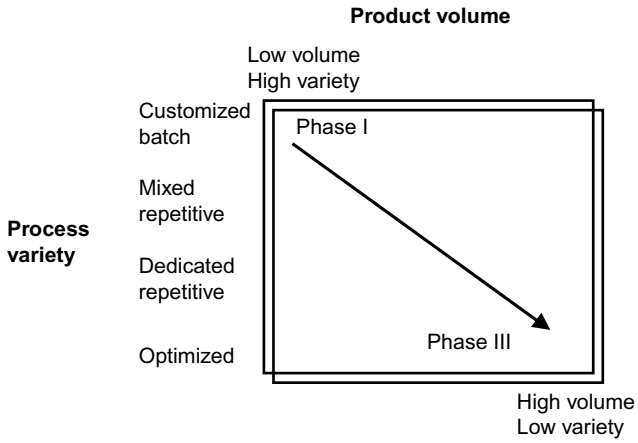


FIGURE 24.3 Relationship of clinical trial phases and process variety and product volume of clinical supplies manufacture.

24.6 CROSS-FUNCTIONAL TRAINING

Cross-functional training and education programs are an important consideration in the clinical trial supplies system. All team members should know the system and their role and contribution toward pushing projects through the finishing line. For example, ensuring that all clinical research associates understand why and how to order study medications and comprehend aspects of lead-time is a simple example of instantly recognizable, value-added education and training. Customer orientation and TQM, which are well known systems in Japanese manufacturing industries, can be used to improve performance within the clinical trial supplies system. Each department should consider how to improve customer satisfaction and the customer relationship. For example, contract manufacturing organizations (CMOs) and third party vendors hired by pharmaceutical manufacturers rely on data regarding clinical trials and supplies to plan and execute their work. Instead of considering these CMOs as suppliers to deliver the clinical supplies for the clinical trials, the CMOs become customers of the pharmaceutical manufacturer whose drug discovery department delivers their bulk active pharmaceutical ingredients (APIs). The bulk APIs and other related clinical trial data should be delivered to CMOs before or at least on schedule to satisfy CMOs. Any interruption in the process, such as a delay in the delivery of the bulk drug substance from pharmaceutical manufacturers, can seriously disrupt the timing and success of their work.

24.7 OUTSOURCING OF MANUFACTURING AND PACKAGING

A drug sponsor may want to contract out its clinical supplies manufacturing and packaging for several reasons. CMOs may provide an expertise or technology required for the project.[6] In addition, contracting out clinical supplies converts the fixed costs of maintaining the personnel and production facilities necessary for manufacturing and packaging the clinical supplies into variable costs of paying CMOs to perform the task and eliminates the need for a set number of people, facilities, and equipment in house. However, many drug sponsors choose to manufacture and package their clinical supplies in-house because they prefer the control and flexibility, as well as the in-house expertise. The question of who will benefit from outsourcing to a CMO might be addressed as why the sponsor cannot or should not do it in-house. The answer is mainly dependent upon just a few factors, such as how a drug sponsor and its board balance risk; whether its product pipeline is full; whether it sets for itself to manufacture and package the commercialized product after NDA approval. There are both advantages and disadvantages for contracting clinical trial supplies, as shown in Table 24.3.

TABLE 24.3 The Advantages and Disadvantages of Contracting Versus In-House Production of Clinical Supplies Manufacturing and Packaging

Advantages	Disadvantages
In-house	*In-house*
1. Maintain oversight of quality—reducing risks of production failures	1. High capital costs because facility/equipment may need to be upgraded or acquired
2. More direct control of outcome—reducing risks of bottlenecks	2. May delay another important project
3. Easy to set or change priorities on multiple projects	
4. Lower cost	
5. Easier for technology transfer process	
6. Gain expertise of manufacturing and packaging the new product	
7. Retain intellectual property	
Contract out	*Contract out*
1. CMOs have expertise and experiences with many companies	1. Resources are needed to manage the CMOs
2. Convert fixed costs into variable costs	2. Difference in culture of the two companies can take more time to work through
3. Save time and cost to upgrade or improve facility/equipment	3. Communication can be more cumbersome
4. Better system in specialized field	4. Less control of project's priority and schedule
5. Can test robustness of the process	5. Confidentiality breach

In general, small companies (especially small biotech companies) have fewer options and are most likely outsource. Their limited development pipelines, lack of capacity, and skill deficiency make the decision clear; the risk of investing in capital and experienced staff outweighs the benefit of in-house cGMP pilot or production plants. Mid-tier companies may have the capability to conduct the cGMP pilot or production plants in-house, but they and their investors must carefully assess the risks of product failure against their ability to survive such an event. Large pharmaceutical companies can build the capacity in-house, but they also see contract manufacturing as a way to reduce risk in their global manufacturing strategy.

Today, dedicated contractors can provide clinical trial manufacturing services. There are more CMOs offering a wide variety of competencies, reliable service, and competitive high quality. As a result, an increasing number of virtual biotech and large pharmaceutical companies are positioned to employ CMOs.

24.8 NEW TECHNOLOGY

The study of average cost for new drug development between 1995 and 2007 was released in Nov. 2014 by the Tufts Center for the Study of Drug Development (CSDD) at Tufts University, in Boston, MA.[7] The study indicated that the average cost of bringing a new drug to the market increased to $2.6 billion compared with the $802 million stated in the previous study released from the same center in 2003.[8] The study mentioned that costs have risen by 145% due to the larger and more complex clinical trials, a greater focus on chronic and degenerative diseases, and higher failure rates. There is no doubt that the clinical trials are more complex; involve more patients, more sites in different countries, more country-specific regulations, and more uncertainty. More and more companies choose to conduct clinical trials in the countries outside the United States in order to lower the trial expenses and increase patient recruitment. Many companies have recently begun to include the direct to patient (DTP) distribution method as an option for patients participating in clinical trials in order to provide more convenience for the patients. The DTP approach delivers clinical trial materials directly to patient's home and picks up biological specimens from the patient home for testing. This DTP approach not only improves patient's willingness and ability to participate in clinical trials, but also increases patient retention throughout the studies, which are the two biggest challenges in clinical trials. Many patients

are unable or unwilling to travel to the clinical trial site for medication refills and testing for many reasons, such as a lack of transportation, patient's deteriorated health condition, and inconvenient location of the clinical trial centers.

Therefore, delivery of the medication to the patient's home can increase patient recruitment and retention. The DPT approach provides advantages from both compliance and patient retention perspectives. However, it is more complex and difficult to control supply logistics. Companies need to ensure that clinical trial materials are properly packaged and delivered to the patient on time and within specifications. Other examples of complex issues affecting clinical supply chains include, but are not limited to, costly comparator drugs, complex clinical trial designs, intricate protocols, delicate investigative compounds and shelf-life, variable dosing schemes, country-specific regulation requirements, complex labeling and packaging approvals, temperature-controlled distribution, and costly shipping across the globe.

A well organized clinical trial planning together with regulatory knowledge and strategy will help companies avoid overproduction and still have enough overage of clinical materials to meet inventory expectation and changes, and be able to deliver the supply when and where it is needed. Timely and accurate forecasting of the clinical trial demand, as well as supply planning and execution, can avoid supply delay, which is very crucial for cost effectiveness and successful results of clinical trials. To prevent clinical trial supply shortage or delay, many companies integrate strategies and information technology to ensure smooth and secured supply of clinical trial materials to the location in a timely and accurate manner.

The use of the new technology and software for full traceability and tracking of supply throughout the life cycle will enable the company to be as efficient as possible. The traceability and tracking starts from order of the clinical material supply, manufacturing of API, randomization, packaging and labeling, release and distribution, until returns and destruction. The software allows a company to manage all clinical trial supply information and inventory from onset throughout its lifetime. In addition, it also increases the visibility of all material delivery and inventory maintained at the warehouse and depot. The traceability and tracking should be not only for investigational product, but also for concomitant medication and all ancillary materials throughout the life cycle. The full traceability and tracking can result in smooth and effective delivery of clinical supply at the lowest possible cost. This planning tool can help the clinical supply manager foresee and solve problems due to the trial design, regulatory, and operational complexities.

Several clinical supply logistic management software are currently commercial available. It is recommended that this type of software should feature 21 CFR-part 11[9] to be in compliance with the current US regulations. 21 CFR-part 11 requires system validation and security of all clinical data software used in conducting clinical trials. Other important features include, but are not limited to, accessibility from any location at all times, barcode/radio frequency identification scanning including 2D and 3D support, demand forecasting, document/report management, including electronic batch record, interfacing and integrating with multiple systems, inventory control management; including lot specific inventory control and tracking throughout the entire supply chain, operations-processing, and supply planning, labeling, and packaging management; including electronic routing of labeling and graphics design and approval, label generation and printing, and label translation to foreign languages, randomization capabilities to ensure accurate assembly of each patient kit per controlled randomization, shipping management; including control of shipment order, processing and returns, supplier and CRO management, supply chain collaboration, supply chain planning, transportation management, and warehouse management. For example, when study designs and dosing requirements are entered, the software can calculate the amount of kits, the dosage form, and all required active ingredients. The software can also analyze different study scenarios with different dosing and/or enrollment patterns to see the impacts on the study supply. This clinical supply management software can not only help the clinical supply manager, but also the warehouse and depot manager, CROs, and quality and regulatory personnel to control, track, and view the activities and status of investigational product from start to finish.

Confusion could easily occur because the clinical trial supply management involves many departments and outside CROs. For example, a CRO outside the country mistakenly destroyed a trial batch because they thought it was no longer needed. Then, when the trial material was needed, the company had to manufacture a new batch as quickly as possible, which could result in trial delay. With the use of clinical supply management software, these types of mistakes could be prevented. The CROs can directly interface with the clinical supply management software to know the status of the batch in real-time, and would not accidentally destroy the batch.

Integrated clinical trial supply management software solution and technologies are currently available to help pharmaceutical industry meet more stringent and complex compliance challenges. For example, clinical investigators in multiple countries worldwide can access a database at the same time using complex

planning tools, which allow clinical trials at multiple centers in many countries to view and track the clinical supply and coordinate and plan more effectively. Similar technologies are also available for clinical manufacturing and packaging groups to electronically view the supply orders in real-time, and effectively collaborate with clinical investigators, instead of using a paper-based system. New software and technologies allow pharmaceutical companies to more effectively address challenges, and improve their new drug development process.

On Nov. 27, 2013, the Drug Supply Chain Security Act (DSCSA) was enacted. The regulation intends to protect public safety by strengthening the security of the drug supply chain, improving detection and removal of counterfeit drug products from the distribution supply chain. The Act outlines a 10-year plan to build an electronic, interoperable system to both identify and trace prescription drugs through distribution across the United States. DSCSA[10–12] requires manufacturers, repackagers, wholesale drug distributors, third-party logistic providers, and dispensers to provide product-tracing information, including three pieces of information: transaction information, transaction history, and transaction statement to their trading partners along the supply chain. The DSCSA currently defines these prescription drugs as finished drugs intended for human use. As the DSCSA implementation plan continues, the FDA will develop and release more guidance during the next few years. Companies should closely follow the regulations and remain in compliance with additions and changes of the regulations as they are released. Comparators used in clinical trials must follow this act.

24.9 GMP CONSIDERATIONS ON MANUFACTURING CLINICAL SUPPLIES

24.9.1 cGMP considerations

The FDA and the pharmaceutical industry have recognized that most early-phase clinical studies are conducted long before the details of product formulation, manufacturing specifications, and production processes are fully defined and validated. Therefore, some of the cGMP requirements are simply not logically applicable to Phase I clinical studies. The FDA expressed a clear recognition of the difference between manufacturing investigational products and commercial products. The agency, however, has a firm commitment to enforce the cGMP requirements and ensure compliance throughout all phases of development, as stated in the 1991 Guideline on the Preparation of Investigational New Drug Products.[13] In Jan. 2006, the FDA published a draft Guidance for Industry: INDs—Approaches to Complying with cGMP's During Phase 1.[14] The FDA is working toward the goal of advancing public health by improving the product development process and increasing the speed of the approval process for new and better medical treatments. The draft guidance reflects the FDA's current thinking regarding quality systems and risk management (or hazard control) principles, which is consistent with the FDA's ongoing cGMP for the 21st century initiative.[15]

The Food, Drug, and Cosmetic Act stipulates that the FDA may approve a NDA only if the methods used in, and the facilities and controls used for, the manufacture, processing, packing, and testing of the drug are found adequate to ensure and preserve its identity, strength, quality, and purity. The FDA's role in the preapproval process is to review data submitted, evaluate, and approve specifications for the manufacture and control of the drug products based on the submitted data, assure cGMP compliance and verify the authenticity and accuracy of the data contained in the NDA application.[16]

The cGMPs are not directly concerned with business performance. The FDA's evaluation is usually directed toward compliance issues such as rejects, reworks, and complaints as indicators of inadequately validated production processes and process deviations as the examples of system noncompliance. In fact, the FDA inspection team frequently visits the reject area early in an inspection to identify failures. They can then go back into the respective batch records to evaluate the cause and to see if adequate remediation has been established and implemented.

FDA inspection teams, highly skilled in drug manufacturing, packaging, and analytical technology, may consist of investigators, analysts, engineers, and/or computer experts, as appropriate. Prior to NDA approval, FDA investigators or inspection teams may perform inspections/data audits, and will validate/verify the proposed methods.[16] The FDA can conduct preapproval inspections or audit any processes involved in the NDA submission including, but not limited to, manufacturing of drug substance(s), biobatch manufacturing, raw materials controls, composition and formulation of finished dosage form, container/closure systems, laboratory support of methods validation, product and process controls, product specifications and test methods, product stability, comparison of relevant preapproval batches and proposed commercial batches, facility-personnel-equipment qualifications, equipment specifications, packaging and labeling controls, process validations, reprocessing, and ancillary facilities (eg, contract testing laboratories, contract packagers). For example, the FDA may audit biobatch manufacturing batch record to determine if

the establishment is in compliance with cGMP requirements. This may include a data audit of the specific batches upon which the application is based (eg, pivotal clinical studies, bioavailability, bioequivalence, or stability batches).

The production of investigational drug products involves added complexity compared with marketed or commercial products due to a lack of fixed routines, variety of clinical trial designs, highly specific packaging designs, and needs for randomization and blinding that increase the risk of product mix up. It is clear that there may be added risk to clinical trial subjects compared with patients treated with commercial products. Compliance to cGMP or QA in manufacture of investigational drug product will minimize the risk to the clinical trial subjects. General cGMP considerations on manufacturing of clinical supplies include:

- Quality management: The investigational clinical trial products or investigational medicinal products (IMPs) must be manufactured under controlled processes per cGMP requirements, and the release of materials should not occur until they meet the predetermined specifications and are approved by the quality unit. This ensures that the investigational clinical trial products used for any trial centers have the same or equivalent quality, purity and safety. The cGMPs are enforced to ensure that there should not be any unsatisfactory manufacturing and packaging processes, which might increase the risk of substandard investigational clinical trial products given to human subjects during clinical studies. The clinical supplies manufacturer must have a qualified person to verify that the investigational clinical trial products manufactured, packaged, repackaged, and imported conform to the product specifications prior to releasing of the batches.
- Manufacturing and process control: In many cases, clinical trial runs of a new drug are produced in facilities other than those used for full-scale production. The facilities and controls used for the manufacture of investigational clinical trial products, especially the NDA biobatches, should conform with cGMPs regulations. Accurate documentation of batch records is essential so that the production process can be defined and related to batches used for early clinical, bioavailability, or bioequivalence studies of a new drug. Specifications, formulations, and manufacturing processes should be as completely defined as possible based on the current state of knowledge. Critical process parameters should be identified, and the respective in-process controls should be implemented. Specifications, formulations, and manufacturing processes should be periodically reassessed during development and updated as necessary. There should be sufficient manufacturing and process controls to ensure the consistency and reproducibility between clinical trial batches, bio/stability batches used in the same or different clinical trials, and a proposed commercial production process. Any changes from initial manufacturing procedures should be fully documented and carried out according to written procedures, which should justify and address any implications for product quality, purity, and safety.
- Documentation: The information and knowledge gained from pharmaceutical development studies provides scientific understanding to support the establishment of specifications and manufacturing controls. Therefore, it is essential that all information and knowledge be fully and properly documented. For example, the acceptance/rejection criteria may not be as specific at early stages; however, these criteria will become more specific and uniform as additional data become available. It is vital that the criteria used and scientific supportive data be fully documented at all stages. The product specifications file (eg, specifications and analytical methods, manufacturing processes, in-process testing and methods, stability data, and storage and shipment conditions) should be continually updated and documented as development of the product proceeds, ensuring traceability to the previous versions.
- Batch records: To prevent any potential mix up or ambiguity, the order in writing for the number of units and/or shipping of the clinical trial product should be received and authorized prior to manufacturing, packaging, and/or shipping of the clinical trial batches. The manufacturing batch records for an investigational drug product must be retained for at least 5 years[17] after the conclusion of the trial or formal discontinuation of the last clinical trial in which the batch was used. During that period, the record must be complete and readily available for inspection.
- Laboratory control: Laboratory equipment and procedures must be qualified and validated. At the early phase of clinical trials, analytical methods performed to evaluate the batches of API used for clinical trials may not yet be fully validated, however there must be enough control based on scientific information. FDA inspectors will review the authenticity and accuracy of data used in the development of a test method.
- Incoming materials for clinical supplies: When more than one supplier and source of each incoming material (especially the active ingredients) was used

during clinical trials, the sponsor should demonstrate that the clinical batches produced from different suppliers and sources of incoming materials are equivalent in terms of conformance with established specifications, including those stated in the NDA application. The sponsor should undertake due diligence to ensure that the incoming materials meet specifications and are produced under a state of control.

- Building and facilities: An addition of any new drug to a production environment must be carefully evaluated regarding its impact on other products already under production and changes that will be necessary to the building and facilities. Construction of new walls, installation of new equipment, and other significant changes must be evaluated for their impact on overall compliance with cGMP requirements. The toxicity, potency, and sensitizing potential may not be fully understood for investigational clinical trial products, which reinforces the need to minimize all risks of cross-contamination. The risk of cross-contamination can be minimized by using properly designed equipment and facilities as well as by having in place written comprehensive methods for manufacturing, packaging, inspecting, sampling, testing, cleaning and sanitizing procedures.[14,18]
- Equipment system: New products, particularly potent drug products, can present cleaning problems in existing equipment. There must be written procedures for sanitation, calibration, and maintenance of equipment, and specific instructions for the use of the equipment and procedures used to manufacture the drug product.[14,18] To prevent the risk of cross-contamination, many companies utilize isolators as well as single use or dedicated equipment during the production of investigational clinical trial products.
- Training program: All personnel involved with investigational products should be appropriately trained in the requirements specific to their duties. Trained, qualified, and experienced persons should be responsible for ensuring that there are systems in place that meet the cGMP requirements. Any training should be documented and kept on file.
- Specifications: Specifications (for starting materials, primary packaging materials, intermediate, bulk products, and finished products), manufacturing formulas, and processing and packaging instructions should be as comprehensive as possible given the current state of knowledge. They should be periodically reaccessed during development and updated as necessary. Each new version should take into account the latest data, current technology used, regulatory and pharmacopoeial requirements, and should allow traceability to the previous document. Any changes should be carried out according to a written procedure, which should address any implications for product quality, such as stability and bioequivalence. "The controls used in the manufacture of APIs for use in clinical trials should be consistent with the stage of development of the drug product incorporating these APIs. Process and test procedure change management should be flexible to provide for changes as knowledge of the process increases and clinical testing of a drug product progresses from preclinical stages through clinical stages."[18]

The FDA recognizes that the experimental nature of the drug substance, formulation, and dosage form at an early stage of development has an impact on establishing specifications. At early stages, the acceptance/rejection criteria may not be as specific; however, it is vital that such criteria be scientifically sound and based upon available scientific data. Specification used as the basis for approval or rejection of components, containers, and closures will become more specific and uniform as additional data become available.[19] The specifications for raw materials and excipients, in-process testing, process evaluation or validation, finished product testing and stability testing for IMPs should be set to assure that the appropriate quality of the product will be maintained.

- Stability: For Phase I clinical trials, it should be confirmed that an ongoing stability program will be carried out with the relevant batch(es) and that, prior to the start of the clinical trial, at least studies under accelerated and long-term storage conditions will have been initiated. When available, the results from these studies should be summarized in a tabulated form. Any supportive data from development studies should be summarized in a tabular overview. An evaluation of the available stability data related to storage conditions stated on the product labels and justification of the proposed period of use, expiration date or retest date assigned to the IMP in the clinical study should be provided.
- For Phase II and Phase III clinical trials, the available stability data should be presented to regulatory agents in a tabulated form. An evaluation of the available data and justification of the proposed shelf-life to be assigned to the IMP in the clinical study should be provided. Data should include results from studies under accelerated and long-term storage conditions.[20]
- Validation: Noncritical manufacturing processes for clinical trial products are not expected to be validated to the extent necessary for routine productions. However, there must be the

combination of controls, calibration, and equipment qualification to ensure that the manufacturing processes for clinical trial products are of the same standard as for commercial products.[21]

- Packaging and labeling control: Clinical trial supplies are normally packaged in an individual container for each subject in the clinical investigation. The packaging and labeling of investigational clinical trial products are usually more complex than commercial drug products, especially when placebo and active products are packaged in the same appearance and usage label for double-blinded clinical studies. Sufficient reconciliations and proper controls should take place to ensure the correct quantity of each product required has been accounted for at each stage of processing. Poor labeling control and accountability may have an adverse impact on the firm's ability to assure that the new drug will always be properly labeled.

> During packaging of investigational medicinal products, it may be necessary to handle different products on the same packaging line at the time. The risk of product mix up must be minimize by using appropriate procedures and/or, specialized equipment as appropriate and relevant staff training.[22]
>
> Packaging and labeling of investigation medicinal products are likely to be more complex and more liable to errors (which are also harder to detect) than for marketed products, particularly when blinded products with similar appearance are used. Precautions against mislabeling such as label reconciliation, line clearance, in-process control checks by appropriated trained staff should accordingly be intensified.[22]

- Distribution of clinical trial material: Distribution of clinical trial material to the trial sites must be in accordance with the local laws and regulations specific to the countries of destination. For example, many countries require clinical trial material labels to be translated into the local language. The clinical trial material must be secured and stored under conditions specified in the label.[17]

In a blinded study, a written emergency unblinding method must be established, trained on, and implemented. "De-coding arrangements should be available to the appropriate responsible personnel before investigational medicinal products are shipped to the investigator site."[22]

The investigator is responsible only for drug accountability at the site and the sponsor is responsible for overall clinical trial materials distributed for use in a trial. Clinical trial material is reconciled at the site level as well as country/affiliate level. Any discrepancies must be documented and investigated. "The delivered, used and recovered quantities of product should be recorded, reconciled and verified by or on behalf of the sponsor for each trial site and each trial period. Destruction of unused investigation medicinal products should be carried out for a given trial site or a given trial period only after any discrepancies have been investigated and satisfactorily explained and the reconciliation has been accepted. Recording of destruction operations should be carried out in such a manner that all operations may be accounted for. The records should be kept by the sponsor."[22]

> When destruction of investigation medicinal products takes place, a dated certificate of, or receipt for destruction, should be provided to the sponsor. These documents should clearly identify, or allow traceability to, the batches and/or patient numbers involved and the actual quantities destroyed.[22]

24.9.2 A risk-based approach

The FDA, the pharmaceutical industry, health care professionals, and patients share a common goal of reliable, high quality, safe, and effective products. There is a need for new products to improve public health. On the other hand, the knowledge about products and processes may not be fully understood for investigational clinical trial products, especially those used in the early phases of clinical study. Through manufacturing science, pharmaceutical firms accumulate knowledge about products, processes, and technology used to manufacture and control these processes, which in turn provide the foundation for a robust quality system. The quality system focuses on critical pharmaceutical quality attributes including chemistry, pharmaceutical formulation, manufacturing process, product quality and performance, and their relevance to safety and efficacy. To produce reliable, high quality, safe, and effective products including early phase IMPs, the desired state of manufacturing is:

- Product quality and performance are achieved and assured through the design of an effective and efficient manufacturing process.[23] Quality should be built into the product instead of relying on product testing alone to ensure product quality. Pharmaceutical firms must continue to move from a compliance mindset to quality by design (building in quality by defining design spaces of product and processes from the development phase and throughout the product's life cycle). The specifications and process parameters are determined and established by the technical experts who have a thorough understanding of

pharmaceutical science, equipment, facilities, and processes. Any variations in materials and processes that can affect the finished product's identity, purity, and quality should be considered during establishment of the specifications and process parameters.

- Innovative and new technologies are used to continuously improve manufacturing and QA. For example, a common risk associated with a highly potent investigational clinical trial material is cross-contamination during manufacturing of the product. Many pharmaceutical firms choose to use single-use disposable equipment and components such as disposable bags, tubing, filters, and storage containers to eliminate cross-contamination risk and reduce the time spent cleaning and setting up equipment. Proper process design, quality assurance, and component selection are important for the successful implementation of single-use disposable equipment and components. PAT is another example of an advanced technology that can facilitate building the knowledge base and mitigate risk. PAT can be used to aid designing, analyzing, and controlling manufacturing processes. PAT allows a better understanding of processes through timely measurement and can be used to quickly evaluate changes that impact product quality, thereby reducing risk during manufacturing processes. PAT changes the traditional concepts of process validation and leads to continuous quality verification strategies.[24]
- Manufacturing science is applied to understand process capability, to assure that the manufacturing processes are preformed in a reproducible manner, and to manage the risk impacting the processes or product quality throughout their life cycle. Elements of risk should be evaluated based on the intended use of a product, patient safety, and the availability of drug product supplies. Continuous improvement efforts must be employed to increase process capability and reduce risk.
- Knowledge is shared between pharmaceutical firms and the FDA. The firms should share sufficient knowledge to provide the FDA with process understanding and rationale for the development of specifications and analytical methods. Adequate pharmaceutical development information, including identification of critical quality parameters and attributes, as well as the justified formulations should be provided to aid in understanding the product and processes.
- Risk management review teams and framework are used by pharmaceutical firms. Pharmaceutical firms should prioritize activities or actions based on a risk assessment, including both the probability of the occurrence of harm and of the severity of that harm. Based on the level of risk, the firm should apply appropriate manufacturing science and develop control strategies in order to prevent or mitigate the risk of producing a poor quality product. The firm's risk-based decisions must be made to assure the identity, purity, potency, and quality regarding the safety and efficacy of new drugs throughout their life cycle (from IND to post marketing phases).

The desired state of manufacturing discussed previously is an objective for the pharmaceutical cGMPs for the 21st century: a risk-based approach, which was announced by the FDA in Aug. 2002 (a risk-based approach aims to ensure resources to be focused on high-risk areas).[25] However, it is the firms' responsibility to ensure that low- and medium-risk areas remain in appropriate states of control because these lower risk classes will receive less FDA regulatory attention. The FDA has expressed its commitment to modernize the pharmaceutical GMP regulations and to streamline the clinical development process to ensure that basic scientific discoveries can be translated more rapidly into new and better medical treatment. In Sep. 2006, the FDA published Guidance for Industry: Quality Systems Approach to Pharmaceutical cGMP Regulations.[23] The guidance, which was intended to serve as a bridge between the 1978 regulations[3] and the FDA's current understanding of quality systems, explains how implementing quality systems can help drug product manufacturers achieve compliance with current GMP requirements. Although the guidance applies to pharmaceutical finished-product manufacturers, it can also be used to guide the clinical supply manufacturing.

24.10 CLEANING VALIDATION AND VERIFICATION

The Code of Federal Regulations PART 211—Current Good Manufacturing Practice for Finished Pharmaceuticals—Equipment Sec. 211.67 on equipment cleaning and maintenance states that "Equipment and utensils shall be cleaned, maintained, and sanitized at appropriate intervals to prevent malfunctions or contamination that would alter the safety, identity, strength, quality, or purity of the drug product beyond the official or other established requirements" and that "written procedures shall be established and followed for cleaning and maintenance of equipment, including utensils, used in the manufacture, processing, packing, or holding of a drug product."[26] In accordance with 21 CFR 211.67, the International Conference on Harmonisation (ICH) has issued recommendations on equipment maintenance and cleaning (Q7A, section

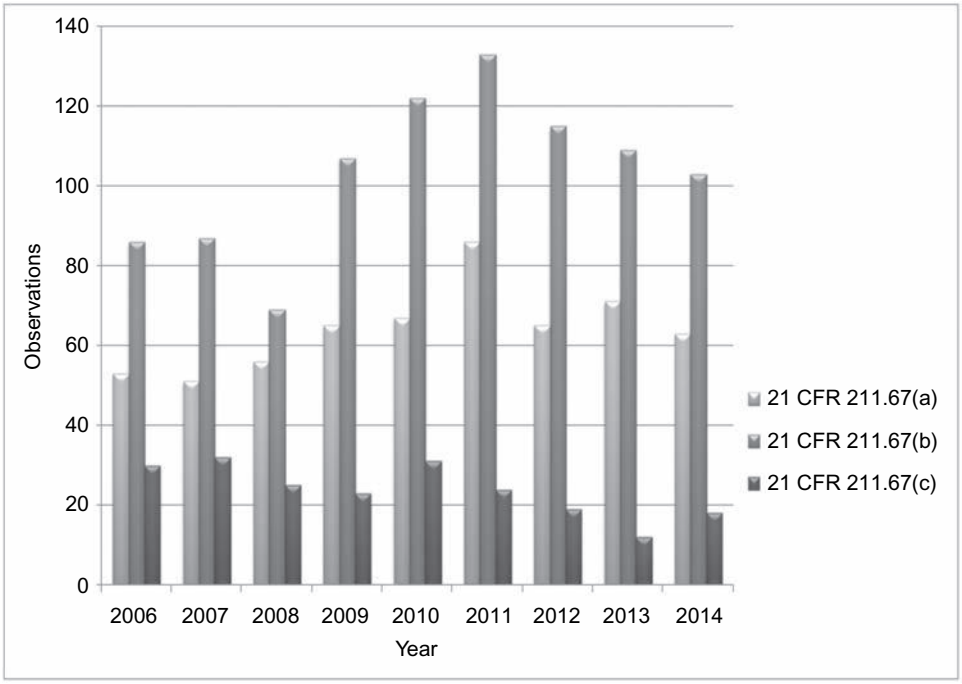


FIGURE 24.4 Frequency of FDA 483 warning letters issued for cleaning violations (21 CFR 211.67) from 2006 through 2014. The light gray columns represent violations of 211.67(a), the medium gray columns represent violations of 211.67(b), and the dark gray columns represent violations of 211.67(c).

5.20–5.26) for compliance and safety that include very similar requirements with more elaboration on specific details.[27] Again, there is a reiteration that written procedures must be established and that detailed cleaning agent selection and preparation, responsibilities, schedules, and cleaning acceptance limits must be documented with rationales. The purpose is to prevent cross contamination between different manufactured lots. This cross-contamination concern applies not only to the API, but also to residual cleaning solvents and detergents.

Even with the seemingly basic requirements outlined, there have been multiple 483 observations issued to pharmaceutical manufacturing facilities since 2006 due to violations of 21 CFR 211.67. 483 observations indicate that a FDA-regulated firm may be in violation of the Code of Federal Regulations (CFR). Fig. 24.4 illustrates the frequency of 483s issued where an infraction against 21 CFR 211.67 was cited (obtained from the FDA website); as stated on the website, this list may not include all of the 483 observations issued.[28] Within 21 CFR 211.67, part a infractions involve satisfying the requirement that "equipment and utensils shall be cleaned, maintained, and, as appropriate for the nature of the drug, sanitized and/or sterilized at appropriate intervals to prevent malfunctions or contamination that would alter the safety, identity, strength, quality, or purity of the drug product beyond the official or other established requirements." Part b infractions require that "written procedures shall be established and followed for cleaning and maintenance of equipment, including utensils, used in the manufacture, processing, packing, or holding of a drug product." Finally, Part c mandates that "records shall be kept of maintenance, cleaning, sanitizing, and inspection as specified in 211.180 and 211.182." Fig. 24.4 illustrates that there may be some ambiguity in compliance with the requirements of 21 CFR 211.67 based upon the number of observations per year, which was approximately 191 from 2006 to 2014, and peaked during this timeframe at 243 total (21 CFR 211.67 parts a–c) observation in 2011.

The ICH guideline on risk management outlines multiple approaches to make and document risk-based decisions.[25] It clearly outlines that the primary goal of risk management is to base the approach on scientific knowledge and to evaluate the impact on the patient. In addition, the level of effort, formality (eg, use of tools), and documentation of the quality risk-management process should be commensurate with the level of risk. The 1993 guidance on cleaning inspections states that for a swab method, recovery should be established from the surface.[29] There are no specific requirements on how to establish these recovery estimates or the acceptance limits, for that matter. It is the manufacturer's prerogative to document the cleaning

rationale (process and acceptance limits) that maintains the quality and purity of the subsequent drug product being manufactured. Therefore, by definition, cleaning validation and verification are areas built upon the premise of risk management.

The purpose of this section is to review the approaches for cleaning equipment used for product manufacturing. The difference between cleaning validation and cleaning verification is outlined first. Acceptance limit calculations are provided that represent either 1/1000 or 10 parts per million (ppm) product carryover in clinical trial operations. The analytical aspects of cleaning method developments are discussed from swab selection, surface selections, to the analytical methodologies that are used for both selective and nonselective approaches. In addition, a guideline for validating a cleaning verification assay is presented.

24.10.1 Cleaning validation versus cleaning verification

Cleaning validation is an extensive, multi-functional program where the entire manufacturing process is considered, ranging from the equipment that will be used to manufacture the products to the analytical methods that will be used to evaluate cleaning effectiveness.[30] Cleaning validation starts with the preparation of a validation protocol. The validation protocol would include rationales regarding cleaning agents and procedures,[31] equipment, equipment swabbing locations,[32] safety-based acceptance limits, selection of products used to demonstrate validation,[30] and validated analytical methodology.

When selecting representative products to validate a cleaning program, several approaches have been presented.[30] The most conservative approach would be to validate the cleaning procedure for all compounds manufactured, which is most likely the approach taken for a marketed product manufacturing facility. However, this approach is expensive, resource consuming, and may not be practical for a multi-compound facility used primarily to support of clinical trials. Therefore, manufacturers may choose to validate on a subset of compounds and examine such parameters as solubility, potency (which drives the acceptance limit), and cleanability. The location and number of swabbing locations for both verification and validation activities are dictated by the product contact surface area (eg, larger components may require a larger number of swabs), energy dissipation (eg, roller compaction or tablet pressing regions are more likely to accumulate product), composition of the product contact areas (ie, polycarbonate vs stainless steel), and cleaning difficulty (tight corners, bends, and hard-to-reach places result in hard-to-clean locations that should be swabbed).[32]

Cleaning verification, on the other hand, is a demonstration for each manufacturing run (typically by submission of cleaning swabs to analytical laboratories) that the API has been removed to a level below the pre-established acceptance limit.

24.10.2 Swab test acceptance criteria

There have been recent proposals that utilize the threshold of toxicological concern (TTC) to calculate an acceptable residue limit.[33,34] The TTC estimates a level of exposure that will not present an appreciable risk to human health. This concept is commonly applied to impurities in drug substances and drug products and a guidance has been recently published (ICH M7).[35] In addition, the European Medicines Agency (EMA) published a guideline outlining the expectation to develop a permissible daily exposure of residuals between veterinary and human products, but the guideline also applies the TTC concept when the residual is a genotoxic API.[36]

When determining the acceptance limit, a strong scientific rationale with relevant factors generally includes: Evaluation of the therapeutic dose of the actives; toxicity of the potential contaminant; solubility of the potential residue; difficulty of cleaning; visual examination; the batch size of the subsequent products on the same equipment; and how the product will be used.[37] Acceptance limit calculations take on the general form where two methods are typically used. Although some authors recommend adjusting the safety factor based upon the dosage form,[37] it is most common to utilize the 1/1000th minimum dose and 10 ppm methods.[38–42] The 1/1000th method assumes that pharmaceuticals are often considered to be nonactive at 0.1% of their prescribed therapeutic dosages.[42] The following equation demonstrates how an acceptance limit is calculated utilizing the 1/1000th method:

Equation (24.1): 1/1000 safety limit calculation

$$\text{Acceptance Limit}\left(\frac{\mu g}{\text{swab}}\right) = \frac{0.001 \times \text{Smallest Strength Product A(mg/day)}}{\text{Maximum Daily \# of Dosage Units Product B(units/day)}} \times \frac{\text{Lot Size Product B(\# of dosage units)}}{\text{Shared Surface Area(cm}^2\text{or in}^2\text{)}} \times \frac{\text{Swab area(cm}^2\text{or in}^2\text{)}}{1 \text{ swab}} \times \frac{1000\ \mu g}{1\ mg}$$

Product A is the product made on the equipment to be cleaned, and product B is the product to be

manufactured on the same equipment after it is used and cleaned for product A.

The second method uses the 10 ppm limit historically used to calculate commercial manufacturing limits. This method allows the maximum carryover of product to be calculated using lot sizes and shared equipment surface area. Eq. (24.2) shows the method used to develop the 10 ppm acceptance limits:

Equation (24.2): 10 ppm acceptance limit calculation

$$\text{Acceptance Limit}\left(\frac{\mu g}{\text{swab}}\right) = \frac{10\ \text{mg}}{1\ \text{kg}} \times \frac{\text{Lot Size Product B(kg)}}{\text{Shared Surface Area(cm}^2\text{or in}^2\text{)}} \times \frac{\text{swab area(cm}^2\text{ or in}^2\text{)}}{1\ \text{swab}} \times \frac{1000\ \mu g}{1\ \text{mg}}$$

After the two limits are calculated using Eqs. (24.1) and (24.2) for each product, the limits are compared. The smaller limit calculated for a product becomes the acceptance limit for the cleaning verification for that product. It is obvious from Eq. (24.2) that two variables drive the acceptance limit: the lot size of product B and the shared surface area. For example, if the shared equipment surface area is constant between two products and the lot size for product B is 50 kg as opposed to 100 kg, the acceptance limit would be calculated as one-half of that for the 100 kg lot size.

24.10.3 Swab selection

Although swab sampling is different conceptually from both the impurity and potency assays, the same scientific rationale governs the development of these assays. Many of the references listed in Tables 24.4 and 24.5 outline different validation approaches. Seno outlined validation practices in the Japanese pharmaceutical industry for cleaning verification,[43] and Kirsch outlined an approach for swab method validation that is consistent with ICH guidelines for method development.[44] An important aspect of any cleaning-verification assay begins with swabbing the surface. Swabs are typically constructed with a polyester knit that will not leave behind fibers after swabbing and possess minimal extractable materials.[45] Jenkins et al. did an extensive evaluation of swabbing materials as a function of the residual particles left on the surface after swabbing.[46] Quartz wool was found to give excellent recovery of analyte; however, it left an excessive amount of particles on the surface. Glass-fiber swabs and cotton swabs also suffered from high levels of residual particles left behind on the surface. Knitted polyester demonstrated the best balance between recovery, residual particle, and background levels. Yang et al. presented a guideline for swabbing techniques that discusses swabbing procedures, solvent selection, and recovery as a function of the coupon tested.[47]

24.10.4 Representative surface selection for method validation

The general expectation is that all product contact surfaces must be considered during cleaning verification in order to demonstrate that the equipment is clean; and it is an expectation that a recovery value is established for each product contact surface during method validation (how to establish this recovery will be discussed next). The recovery is used to correct the submitted swab result for incomplete removal of the residual analyte from the surface and then compare that result versus the acceptance limit. For example, if 80% recovery was obtained during method validation, that value would be incorporated into the analytical method. Let's assume that the uncorrected analytical result was 4.2 μg/swab compared with an acceptance limit of 5.0 μg/swab. After correcting for the recovery of 80%, the result is now 5.25 μg/swab, which constitutes a failing result in this example. In this case, the importance of the recovery value becomes apparent because it is the difference between an acceptable result and a failing result.

All product manufacturing product contact surfaces would ideally be represented by the analytical method. However, the number of product contact surfaces and the sheer number of manufacturing equipment types and suppliers makes verifying recovery of each analyte off of each surface quite onerous. For example, equipment in a CT manufacturing area is shared among many compounds in the portfolio, especially for a pipeline that is rich in Ph I compounds. As new equipment is purchased, it may have new product contact surface material of construction (MOC). Each compound in the portfolio that utilizes this piece of equipment would require a method revalidation in order to add a recovery factor for the new product contact surface MOC. As the number of different materials of construction increases along with the number of compounds being manufactured, the difficulty and complexity of sustaining that approach also increases.

However, there are approaches that allow for a subset of the total number of surfaces to be evaluated during method development. Stainless steel is the most common material in a manufacturing environment.[48] Thus, some companies establish recovery for stainless steel and apply that correction to all swab submissions. It could be outlined that if the product contact surface

TABLE 24.4 Nonselective Analytical Techniques for Cleaning Verification With Associated Analytes and Reported/Estimated Limit of Detection (LOD) and Limit of Quantitation (LOQ), if Available

Nonselective techniques				
Detection technique	**Analyte**	**Reported/estimated LOD[a]**	**Reported/estimated LOQ[a]**	**Reference**
pH	Cleaning agents	200 μg/mL	666 μg/mL	52
Conductivity	Cleaning agents	20–200 μg/mL	67–666 μg/mL	52
TOC	Sulfacetamide, sulfabenzamide, sulfathiazole	low μg/cm^2	low μg/cm^2	53
	Total carbon	<50 ppb	167 ppb	51
	Detergents	50 ppb	167 ppb	54
	Cleaning agents	0.95 μg/cm^2	3.2 μg/cm^2	55
	Cleaning agents	20–2000 μg/mL	67–6660 μg/mL	52
	Aspirin	3–15 ppm	10–50 ppm	56
Visual	Residues	62.5 μg/100 cm^2	208 μg/100 cm^2	57
	Cleaning agents	20–2000μg/mL	67–6660 μg/mL	52
	Aspirin	>0.1 μg/cm^2	0.3 μg/cm^2	56

[a]*LOQ estimated assuming relative standard deviation (RSD) = 10% and the LOD estimated assuming RSD = 33%.*

TABLE 24.5 Selective Analytical Techniques for Cleaning Verification With Associated Analytes and Reported/Estimated Limit of Detection (LOD) and Limit of Quantitation (LOW). When Compound Not Indicated, Proprietary Listed as Analyte

Selective techniques				
Detection technique	**Analyte**	**Reported/estimated LOD[a]**	**Reported/estimated LOQ[a]**	**Reference**
HPLC-UV	Amlodipine	0.02 μg/mL	0.7 μg/mL	58
	Ertapenem	0.0006 μg/mL	0.002 μg/mL	59
	Sulfadoxine	0.12 μg/mL	0.4 μg/mL	60
	Acetylsalicylic acid	0.04 μg/mL	0.1 μg/mL	61
	Sumatriptan succinate	0.003 μg/mL	0.01 μg/mL	62
	Losoxantrone	0.005 μg/mL	0.02 μg/mL	63
	Bisnafide	0.004 μg/mL	0.01 μg/mL	64
	Nonoxynol-9	0.06 μg/mL	0.2 μg/mL	65
UPLC-UV	Dienogest, finasteride, gestodene, levonorgestrel, estradiol, ethinylestradiol, and norethisterone acetate	0.02–0.20 μg/mL	0.05–0.35 μg/mL	66
UV	Any compound with chromophore	Typically higher than HPLC-UV	Typically higher than HPLC-UV	No references for cleaning verification
HPLC-MS	Proprietary	0.00002 μg/mL	0.00007 μg/mL	67
	Proprietary	0.0005 μg/mL	0.002 μg/mL	68
	Proprietary	<0.004 μg/mL	0.01 μg/mL	69
Ion mobility spectrometry (IMS)	Diphenylhydramine	0.0009 μg/mL	0.003 μg/mL	70
	Tamoxifen	0.05 μg/mL	0.2 μg/mL	71

(*Continued*)

TABLE 24.5 (Continued)

Selective techniques				
Detection technique	**Analyte**	**Reported/estimated LOD[a]**	**Reported/estimated LOQ[a]**	**Reference**
	Proprietary	0.25 μg/mL	0.8 μg/mL	72
	Duloxetine, citric acid, tetrasodium ethylenediaminetetraacetic acid (EDTA), and a proprietary amphoteric surfactant, CIP 200	<5 μg/swab	5 μg/swab	73
Gas chromatography	methenamine hippurate	Not reported	Not reported	74
HPLC-ELSD	2-amino-bicyclo[3,1,0]hexane-2,6dicarboxylic acid	0.26 μg/mL	0.9 μg/mL	75
TLC	Chloramphennicol	0.003 μg	0.01 μg	76
OPLC	Steroid hormones	0.03 μg	0.1 μg	77
CE	No reference to compounds	Typically higher than HPLC-UV	Typically higher than HPLC-UV	78–80
Atomic absorption	Cisplatin	0.0005 μg/mL	0.002 μg/mL	81
ICP-AES	Lithium	0.033 μg/mL	0.1 μg/mL	82
HPLC-electrochemical	Clarithromycin	0.3 μg/mL	1.0 μg/mL	83
	Isoproterenol sulfate	0.1 ng/mL	0.3 ng/mL	84
Spectroscopy	Proprietary	4 μg/cm^2	13 μg/cm^2	85
	Cleaning agent	Not stated	Not stated	86
	bovine serum albumin	low μg/cm^2	low μg/cm^2	87
HPLC-fluorescence	Norfloxacin	5 ng	20 ng	88
HP-TLC-fluorescence	Norfloxacin	5 ng	20 ng	88
Ion exchange conductivity	Cleaning agent	0.1 μg/mL	0.3 μg/mL	89
	Cleaning agent- CIP 100 determined as EDTA	0.13 ppm	0.39 ppm	90
Ion exchange–UV	Cleaning agent	1 μg/mL	3 μg/mL	91
Charged aerosol detection	Mometasone Furoate	0.6 ng	2 ng	92
	Albuterol	1.5 ng	5 ng	
	Loratadine	1.4 ng	5 ng	
Light-induced fluorescence (LIF)	Proprietary		0.353 μg/mL	93
	Amlodipin	5–50 μg/dm^2	50–300 μg/dm^2	94
	Caffeine			
	Ibuprofen			
	Losartan			
	Nifuroxazid			
	Paracetamol			
Dry-swabbing/image analysis (DSIA)	Rutin			
	Valsartan			
	Endiex			
	Ibalgin			
	Lozap H			
	Valzap			

[a]*LOQ estimated assuming RSD = 10% and the LOD estimated assuming RSD = 33%.*

area is less than, for example, 1.0%, that surface MOC is not a high risk for contamination and does not need to be included in the validation. Another option for MOCs not included in the analytical method validation would be to assume some percentage of recovery seen on stainless steel (eg, 10%), a percentage based upon similar materials, or possibly the lowest percent recovery seen on any evaluated MOC. All of these assumptions should be based upon a strong scientific rationale and supported by internal procedures and processes. Finally, two authors have advocated that all MOCs utilized in the manufacturing environment should be placed into groups.[48,49] At that point, a representative of a group would be included in the analytical method development. The authors advocate that many MOCs behave similarly when validating an analytical method; by choosing representative recovery surfaces for those surfaces other than stainless steel rather than doing recovery experiments on all surfaces, the effort can be focused proportionally on the risk while still maintaining patient safety.

This approach of grouping MOC for analytical method development in support of cleaning verification/validation activities is an excellent opportunity to apply a quality risk management approach.

24.10.5 Analytical methodologies

Cleaning verification methods for swabs are typically validated by demonstrating analyte recovery from a "test coupon." As discussed previously, this test coupon can be the exact MOC in the manufacturing area or it may be a scientifically defined "representative." While analyzing rinse solutions has its place in the manufacture of API, it is not a common practice in support of a solid oral drug product clinical trial manufacturing site. In these analyte recovery experiments, a known amount of an analyte will be spiked onto the coupon; the coupon will then be swabbed in a methodical manner illustrated in Fig. 24.5. For example, 10 swipes are performed in the vertical direction, the plate or test coupon is rotated, and 10 additional swipes are performed in the horizontal direction for a 10 cm × 10 cm coupon. For a larger surface (eg, 25 cm × 25 cm), 20 swipes in each direction may be required. Some researchers have advocated zig-zag patterns rather than horizontal and vertical swipes. Typically, there will be a training program required that teaches the swabber proper technique. It is important to be consistent (eg, the amount of pressure applied to the swab and potentially the angle of the swab head with regard to the surface) and that the validation mimics how the swabs will be collected from the manufacturing equipment. The swab will then be extracted in a solvent that maximizes solubility and is compatible with the detection technique, and ultimately analyzed. Depending upon the analyte, the acceptance limit, and the swab, 5–20 mL is an appropriate volume of extraction solvent to ensure a reproducible extraction without overdiluting the analyte, which could ultimately impact analytical detection.

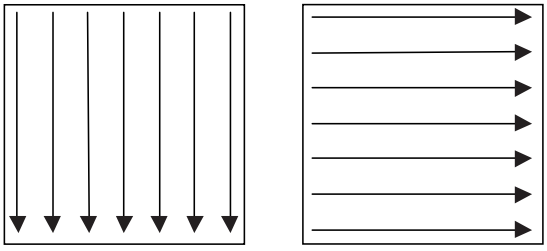


FIGURE 24.5 Illustration of typical swabbing practice. The arrows indicate the swabbing direction on the test coupon.

With the advances in analytical technology, it should be a rare exception that the toxicologically determined cleaning limit cannot be attained from the perspective of determination of the analyte; although it does become more difficult to quantitatively recover low amounts of analytes from surfaces. When selecting the appropriate analytical technique for cleaning verification, several parameters must be considered, such as the safety acceptance limit for the cleaning verification assay, solubility,[50] the molar absorptivity of the molecule, or ionization or oxidation potential if alternative detection techniques are an option. In addition, the benefits of requiring a selective assay versus a nonselective assay should be considered. In a nonselective assay, the blank of a rinse and swab blank are determined to ascertain if there is any deviation from the control (ie, the blank in the same solution). The advantage of this approach includes the fact that a variety of residues may be detected, which include cleaning agents, excipients, previously manufactured products, degradation products, and APIs. It should, however, be realized that the identity of the contaminant would remain unknown after this assay. The analyst would only know that the equipment had a detectable residue, which may or may not help improve the cleaning process on subsequent manufactures. Many nonselective analytical techniques are outlined in Table 24.4. Within this table are indicated the analytes, and the reported limit of detection (LOD) for the technique in association with the cleaning assay. The most prevalent of the nonselective assays is total organic carbon (TOC). It has the advantages of good sensitivity, short analysis time, and being capable of detecting all residual carbon—independent of the source.[51] For a selective technique, the target analyte is known before method validation. Table 24.5 outlines many of the feasible selective analytical techniques.

The list is not all-inclusive because of the multiple detection schemes that can be implemented with high performance liquid chromatography (HPLC). HPLC is commonly used in the pharmaceutical industry. As a result, cleaning verification assays by this approach outnumber all others. Typically, at the time of manufacture, a potency assay has been developed and validated for the drug product. From a conservative standpoint, anything that elutes, or coelutes, at the retention time of the analyte of interest will be attributed to the API. Therefore, the swab method is usually very similar, if not identical, to the potency method. Characteristic swab assays would most often be isocratic methods with reduced run times. In addition to HPLC with ultraviolet (UV) detection, many other detection schemes are outlined in Table 24.5. The analytical chemist chooses these techniques based upon the chemical properties of the target molecule and the cleaning acceptance limit, which was previously established based upon dose and equipment surface-area considerations. For example, a molecule that does not possess a chromophore and is somewhat potent would pose an interesting analytical challenge. The analyst might consider evaporative light scattering detection, electrochemical detection, or mass spectral detection, to name a few possibilities. There are a few examples of analytical techniques listed in Table 24.5 where a specific example for a cleaning verification assay could not be located; however, depending upon the molecular properties of the compound, the cited technique would be a viable option for a cleaning verification assay.

24.10.6 Analytical method validation

Regardless of the analytical technique chosen for cleaning verification, there are certain components that are expected of any method validation activity and outlined in the ICH guidelines on analytical methods Q7A.[27] For a quantitative impurity test, validation is required for specificity, limits of detection, and quantitation, linearity, accuracy, precision (repeatability and intermediate precision, depending upon the phase of development), and range (ICH Q2(R1)).[95] Each of these validation criteria is listed next with a brief explanation on the applicability to cleaning verification assays. A limit test for cleaning verification assays is often employed during clinical trials. With a limit test, a result of "pass" or "fail" is reported. The limit test significantly simplifies the validation of the cleaning verification assay. In order to validate a limit test, specificity, limit of quantitation (LOQ), accuracy, and stability must be evaluated at the acceptance limit. Linearity and range may be excluded. A drawback of the limit test is that the ability to trend the cleaning efficiency (except through failures) for a particular process, or compound, is eliminated because a quantitative result is not reported. Due to the fact that the formulation and process are likely to evolve and the attrition of compounds during development is high, one might argue that trending doesn't provide much value here. However, if the compound makes it to market, a quantitative assay is expected for commercial manufacturing.

24.10.6.1 Specificity

Specificity should be demonstrated using interferences (ie, extractables) that arise from the cleaning swabs and/or blank surfaces. Specificity between the excipients utilized in the drug product manufacture is desirable, but not required. If specificity is not achieved it must be assumed that the peak response is from the API. If the interference is low enough, it may be acceptable as long as it does not unduly bias the recovery results and/or it can be accounted for appropriately through background subtraction. Significant interference from a surface may indicate compatibility issues for the selected sampling solvent and may vary (drop) over time as surfaces on equipment are repeatedly cleaned.

24.10.6.2 Detection and quantitation limits

The LOQ, as opposed to the LOD, is the most important attribute of the swab method because it has been previously determined, from a safety and cross-contamination standpoint, that residual material above the cleaning acceptance limit is of serious concern. The LOQ is defined as a percent relative standard deviation (% RSD) of 10%, and the LOD is defined as a % RSD of 33%. These values are typically determined on low-level injections of the standard. The sensitivity of the method must be low enough to ensure sample responses near the lowest acceptance limit are quantifiable. For typical swab assays, this requirement may necessitate an LOQ of 1 μg/mL (5 μg/swab limit, 1 swab into 5 mL of solvent). When discussing an acceptable precision for a swab assay (including recovery from surfaces), larger % RSD values may be acceptable (eg, 20%). However, the assay pass-fail limit should be lowered so that this variability is accounted for. For example, if 80% recovery was obtained during validation with a 15% RSD for a 5 μg/swab limit, the assay pass-fail limit could be reasonably lowered to 68% to take into account the swabbing variability.

24.10.6.3 Linearity

Linearity is assessed to cover the full range of acceptance limits required by the method. Linearity has been performed in a couple of different manners. First, linearity has been performed through the preparation

of standards that are equal to the resulting swab solution concentration after putting the swab into solvent. The second approach would involve spiking plates at many different levels and evaluating the recovery at each level. The second approach is problematic because recovery is usually observed to decrease as the limit at the spiked level on the surface decreases. Thus, it may be unreasonable to expect linearity over a large range. However, if the range is kept narrow, either approach is suitable.

24.10.6.4 Accuracy and recovery

When compared with a typical potency assay, this validation parameter is the most variable.

Equation (24.3): accuracy/recovery calculation

$$\text{Recovery}(\%) = \frac{\text{spiked amount determined}(\mu g)}{\text{actual spiked amount}(\mu g)} \times 100$$

It is not uncommon to see RSDs of recovery results for a cleaning range from 2% to 10% or higher depending upon the surface. In addition, the recovery of an analyte from the test surfaces is not expected to be complete. Depending upon the surface's characteristics (eg, a smooth polished surface vs a rough surface), acceptance limit (eg, 0.5 mcg/swab vs 100 mcg/swab), and the solubility of the compound, recoveries as low as 20%, or lower, may be acceptable and accounted for in the method calculations through the use of recovery factors. The minimum sample response (extract concentration) must not be less than the quantitation limit of the assay to ensure that analyte levels close to the lowest acceptance limit are quantifiable.

24.10.6.5 Intermediate precision

With the previously mentioned variability that may be observed during accuracy/recovery determinations, it must be realized that acceptance criteria around intermediate precision must be set accordingly. For example, it is quite likely that if Analyst A could obtain a recovery of 80% for a 5 μg/swab limit, Analyst B could obtain a value of only 70% for the recovery. When looking at the residual surface analyte difference between the two analysts, Analyst B would have left only 0.5 μg (0.02 $\mu g/cm^2$ for a 2.5 cm × 2.5 cm test coupon) across the tested surface area. It is easy to see that swabbing technique can play a crucial role in swabbing success when such small amounts are considered. As the compound becomes more potent, the residual differences between Analyst A and Analyst B may be in the nanogram range.

24.10.6.6 Range

The range of acceptable method performance is based upon the LOQ and the linearity assessment.

24.10.6.7 Standard and sample stability

Stability of the unextracted swabs is determined in order to allow time for transport of the swabs between the swabbing site and the analytical testing site. Standard stability and swab extract stability are determined to facilitate analytical testing. Swab and extract stability must be assessed after contact with each of the product contact surface materials because sample stability can be affected by surface-specific contaminants.

24.11 CASE STUDY

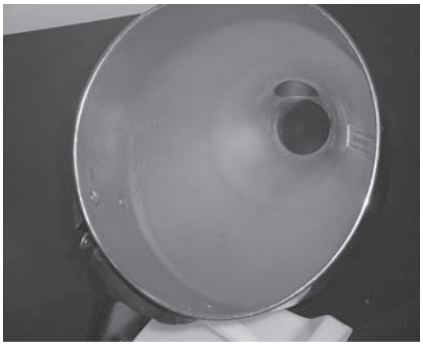
24.11.1 Example

The following case study is for a tablet manufactured on a Manesty tablet press in the clinical trial manufacturing area for a first human dose trial. The compound will be manufactured at doses of 1 mg/tablet, 25 mg/tablet, and 100 mg/tablet and has a good molar absorptivity. The 1/1000th minimum dose and 10 ppm methods were used to calculate the swab test acceptance limits, assuming an equivalent lot and dosage strength will be subsequently manufactured. The analysis shown next outlines the selection of swabbing locations, acceptance limit calculations, and analytical methodology based upon previously mentioned criteria.

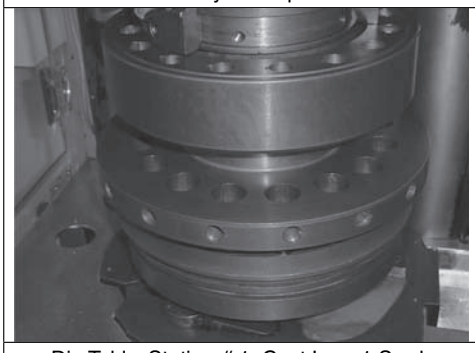
Fig. 24.6 is an illustration of the Manesty press. There are magnified views indicating swab locations. All of this information (picture, number of swabs, surface areas in Table 24.6, and MOC) is part of the cleaning master plan and is generated for each piece of manufacturing equipment and each swab location. This documentation practice serves as a guide for documenting swab locations and rationales for such locations.

Based upon product contact surface area (outlined in Table 24.7) the surface areas were calculated for each component and it was determined that three equipment components had large surface areas. In this case, the discharge port, the die table, and the hopper all had significant surface area contributions. In addition, the die table would be considered a high-energy dissipation area where it is likely that product would accumulate during manufacture. The materials of construction are stainless steel and cast iron; therefore both surface types are swabbed. In turn, an analytical method should be validated where the recovery from stainless steel and cast iron are evaluated. None of the other areas were evaluated as especially difficult to clean in this example. Thus, no additional swabs were added.

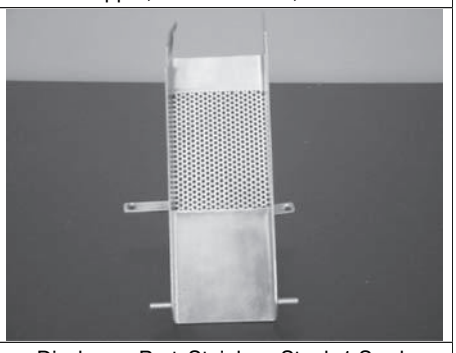
Manesty tablet press



Hopper, stainless steel, 1 swab



Die Table, Station # 1, Cast Iron, 1 Swab



Discharge Port, Stainless Steel, 1 Swab

FIGURE 24.6 Manesty tablet press with swab locations, material of construction, and number of swabs required per location indicated.

TABLE 24.6 Material of Construction and Surface Area Calculations for the Product Contact Surfaces Illustrated Above

Component	Material	Surface area (cm^2)	Surface area (%)	# Swabs
Discharge port	Stainless steel	42.0	5.4	1
Upper punch	Stainless steel	0.6	0.1	0
Lower punch	Stainless steel	1.1	0.1	0
Die	Stainless steel	0.9	0.1	0
Die table, station 1	Cast iron	126.5	16.3	1
Feeder bowl	Cast iron	145.5	18.8	0
Hopper	Stainless steel	458.9	59.2	1

TABLE 24.7 Lot Size and Dose Strengths for Case Study

Product	Lowest dose strength (mg/tablet)	Lot size (kg)	Lot size (units)	Maximum daily dose (units)	Shared surface area (cm^2)
A	1	20	10,000	1	
B	10	1	5000	5	775.5

24.11.2 Acceptance criteria

Using the 1/1000th minimum dose and 10 ppm methods, the following equations list the product and equipment characteristics used to calculate the swab acceptance limits for the case study.

1. 1/1000th Method:

$$\text{Acceptance Limit}\left(\frac{\mu g}{\text{swab}}\right) = \frac{0.001 \times 1 \text{ mg/day}}{5 \text{ units/day Product B}} \times \frac{\text{Product B}(5000 \text{ dosage units})}{\text{Shared Surface Area } 775.5 \text{ cm}^2} \times \frac{\text{Swab area } 25 \text{ cm}^2}{1 \text{ swab}} \times \frac{1000 \ \mu g}{1 \text{ mg}}$$

Acceptance Limit = 32.2 μg/swab

2. 10 ppm Method:

$$\text{Acceptance Limit}\left(\frac{\mu g}{\text{swab}}\right) = \frac{10 \text{ mg}}{1 \text{ kg}} \times \frac{\text{Lot Size Product B } 1 \text{ kg}}{\text{Shared Surface Area } 775.5 \text{ cm}^2} \times \frac{\text{Swab area } 25 \text{ cm}^2}{1 \text{ swab}} \times \frac{1000 \ \mu g}{1 \text{ mg}}$$

Acceptance Limit = 322 μg/swab

In order to keep the case study simple, only one piece of equipment was used to illustrate the selection of swab locations, images included in CT master plan, and calculation of limits. Tablet products are typically made with several steps using several pieces of equipment such as a granulator/dryer, mill, blender, press, or optional coater. The total product contact surface area of all equipment should be used for calculation of limits. The smallest limit calculated for a product using all of the calculations becomes the acceptance limit for the cleaning verification of that product. In this case, the acceptance limit would be 32 μg/swab. If product B, the lot to be manufactured next, is unknown at the time of product A's manufacture, the worst case should be assumed in the calculation. For example, if the smallest lot that has ever been manufactured in the facility is 1000 dosage units or 1 kg, those values should be used in the equation to generate a maximum allowable carryover of product A.

24.12 SUMMARY

This chapter presents a strategy to effectively manage the complex clinical supplies manufacturing process. This strategy includes planning, collaborative communication among the cross-functional team, lean manufacturing, effective training programs, and application of new technologies to facilitate the clinical supplies manufacturing. cGMP considerations to ensure the quality of the clinical materials during the clinical supplies manufacture are also discussed. The comprehensive clinical supply manufacture strategy and cGMP controls prevent clinical supply manufacture from becoming an obstacle to the clinical trial program and consequently reduce the total cycle time of introducing a new drug to market. In addition, various aspects of cleaning verification and considerations for a cleaning validation program are presented. Acceptance limit calculations and analytical methodology and validation were outlined to illustrate the vast interest in the topic of cleaning verification in a clinical trial manufacturing facility.

Acknowledgments

The authors would like to acknowledge the contributions of Lisa Ray, Jack Y. Zheng, and Richard Hwang to the previous version of this chapter.

References

1. FDA Guidance for Clinical Investigators., Institutional Review Boards and Sponsors Process for Handling Referrals to FDA Under 21 CFR 50.54 Additional Safeguards for Children in Clinical Investigations; Draft, May 2006.
2. Merck Seeks a Cure: D.F. Carr and E. Cone, Baseline the Project Management Center, June 8, 2004.
3. Monkhouse DC, Rhodes CT, editors. *Drug products for clinical trials: an international guide to formulation, production, quality control*. New York, NY: Marcel Dekker, Inc; 1998.
4. Paez O, Dewees J, Genaidy A, Tuncel S, Karwowski W, Zurada J. The lean manufacturing enterprise: an emerging sociotechnological system integration. *Hum Factors Ergon Manuf* 2004;**14** (3):285–306.
5. Womack JP, Jones DT. *Lean thinking, banish waste and create wealth in your corporation*. New York, NY: Free Press; 2003.
6. Broeze RJ, Macaloney G, Parker S. Executive roundtable: cutting costs with contract manufacturing, next generation pharmaceutical article; September 2004.
7 DiMasi JA. November 18 *Cost of developing a new drug*. Tufts Center for the Study of Drug Development, Tufts University Boston MA; 2014.
8 DiMasi JA, Hansen RW, Grabowski HG. The price of innovation: new estimates of drug development costs. *J Health Econ* 2003;**22**:151–85.
9 FDA. Guidance for Industry: Part 11, Electronic Records; Electronic Signatures-Scope and Application. Available from: <www.fda.gov/cder/guidance/>.
10 FDA. Guidance for Industry: DSCSA Implementation: Product Tracing Requirements – Compliance Policy. Available from: <www.fda.gov/cder/guidance/>.

11 FDA. Draft Guidance for Industry: Drug Supply Chain Security Act Implementation: Identification of Suspect Product and Notification. Available from: <www.fda.gov/cder/guidance/>.
12 FDA. Draft Guidance for Industry: Drug Supply Chain Security Act Implementation: Identification of Suspect Product and Notification. Available from: <www.fda.gov/cder/guidance/>.
13. FDA Guidance for Industry: Guideline on the Preparation of Investigational New Drug Products (Human and Animal); March 1991.
14. FDA Guidance for Industry: INDs—Approaches to Complying with cGMP's during Phase 1; Draft January 2006.
15. PQRI/FDA Report on the Workshop, A Drug Quality System for the 21st Century, April 22–24, 2006, Washington, DC.
16. FDA Preapproval Inspections (Compliance Program #7346.832).
17. Directive 2003/94/EC Article 13, 2.
18. FDA Guidance for Industry: Q7A Good Manufacturing Practice Guidance for Active Pharmaceutical Ingredients; August 2001.
19. FDA Guidance for Industry: Content and Format of Investigation New Drug Applications (INDs) for Phase 1 Studies of Drugs, Including Well-Characterized, Therapeutic, Biotechnology-Derived Products; November 1995.
20. Good Manufacturing Practice in Pharmaceutical Production, World Health Organization.
21. FDA Policy Guide: Process Validation Requirements for Drug Products and Active Pharmaceutical Ingredients Subject to Pre-Market Approval, CPG 7132c.08; March 2004.
22. EU GMP Annex 13.
23. FDA Guidance for Industry: Quality Systems Approach to Pharmaceutical cGMP Regulations; September 2006.
24. FDA Guidance for Industry: PAT—A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance; September 2004.
25. FDA Guidance for Industry: Q9 Quality Risk Management; June 2006.
26. 21 CFR 211.67, Code of Federal Regulations PART 211—Current Good Manufacturing Practice for Finished Pharmaceuticals—Equipment cleaning and maintenance, Subpart D—Equipment.
27. ICH Q7A, Good Manufacturing Practice Guide for Active Pharmaceutical Ingredients.
28. U.S FDA, Center for Drug Evaluation and Research Website. <http://www.fda.gov/ICECI/Inspections/ucm250720.htm>.
29 Validation of Cleaning Processes (7/93), Guide to Inspections Validation of Cleaning Processes. <http://www.fda.gov/ICECI/Inspections/InspectionGuides/ucm074922.htm>.
30. Garvey CR, Levy S, McLoughlin T. Cleaning Validation–Oral Solid Packaging Equipment. *Pharm Eng* 1998;**18**(4):20–6.
31. Hwang R-C. New Challenges in Cleaning Validation. *Am Pharm Rev* 2002;**5**(3):42–6.
32. Gerber M, Perona D, Ray L. Equipment Cleaning in Clinical Trial Material Manufacturing and Packaging. *Pharm Eng* 2005;**25**(3):46–54.
33 Dolan DG, Naumann BD, Sargent EV, Maier A, Dourson M. Application of the threshold of toxicological concern concept to pharmaceutical manufacturing operations. *Regul Toxicol Pharmacol* 2005;**43**:1–9.
34 Bercu JP, Dolan DG. Application of the threshold of toxicological concern concept when applied to pharmaceutical manufacturing operations intended for short-term clinical trials. *Regul Toxicol Pharmacol* 2013;**65**:162–7.
35 ICH Harmonised Tripartite Guideline: Assessment and Control of DNA Reactive (Mutagenic) Impurities in Pharmaceuticals to Limit Potential Carcinogenic Risk M7 Current Step 4 version dated 23 June 2014.
36 Guideline on setting health based exposure limits for use in risk identification in the manufacture of different medicinal products in shared facilities, 20 November 2014 EMA/CHMP/ CVMP/ SWP/169430/2012 Committee for Medicinal Products for Human Use (CHMP) Committee for Medicinal Products for Veterinary Use (CVMP)
37. Hall WE. Cleaning and Validation of cleaning for Coated Pharmaceutical Products. *Drug Manuf Technol Ser* 1999;**3**:269–98.
38. Agalloco JP., et al. Points to Consider for Cleaning Validation. Technical Report No. 29. PDA J Pharm Sci Technol (USA) 1998; 52: 1–23.
39. LeBlanc DA. Establishing Scientifically Justified Aceptance Criteria for Cleaning Validation of Finished Drug Products. *Pharm Technol* 1998;**22**(10). pp. 136–148.
40. Dolan DG, Naumann BD, Sargent EV, Maier A, Dourson M. Application of the threshold of toxicological concern concept to pharmaceutical manufacturing operations. *Regul Toxicol Pharmacol RTP* 2005;**43**(1):1–9.
41. Zeller AO. Cleaning Validation and Residue Limits: A Contribution to Current Discussions. *Pharm Technol Europe (USA)* 1993;**5**. pp. 18, 20, 22, 24–27.
42. Fourman GL, Mullen MV. Determination of Cleaning Validation Acceptance Limits for Pharmaceutical Manufacturing Operations. *Pharm Techn (USA)* 1993;**17**. pp. 54, 56, 58, 60.
43. Seno S, Ohtake S, Kohno H. Analytical validation in practice at a quality control laboratory in the Japanese pharmaceutical industry. *Accredit Qual Assur* 1997;**2**(3):140–5.
44. Kirsch RB. Validation of analytical methods used in pharmaceutical cleaning assessment and validation. *Pharm Technol* 1998;**22** (Suppl. 2):40–6.
45. Miscioscio K, Cooper DW. Characteristics of wipers and swabs for pharmaceutical applications. *J IEST* 2000;**43**(1):31–7.
46. Jenkins KM, Vanderwielen AJ, Armstrong JA, Leonard LM, Murphy GP, Piros NA. Application of total organic carbon analysis to cleaning validation. *PDA J Pharm Sci Technol* 1996;**50**(1):6–15.
47. Yang P, Burson K, Feder D, Macdonald F. Method development of swab sampling for cleaning validation of a residual active pharmaceutical ingredient. *Pharm Technol* 2005; **29**(1):84–94.
48 Pack B, Hofer J. A Risk-Management Approach to Cleaning-Assay Validation. *Pharm Technol* 2010;**34**(6):48–55.
49 Forsyth RJ, O'Neill JC, Hartman JL. Cleaning Validation: Grouping Materials of Construction Based on Recovery Data. *Pharm Technol* 2007;**31**(10):102–16.
50. PhRMA Quality Committee, Lazar M. PhRMA Guidlines for Validation of Cleaning Procedures for Bulk Pharmaceuticals. *Pharmaceutical Technol* 1997;56–73.
51. Wallace B, Stevens R, Purcell M. Implementing Total Organic Carbon Analysis for Cleaning Validation. *Pharm Techn Aseptic Process* 2004;40–3.
52. Westman L, Karlsson G. Methods for detecting residues of cleaning agents during cleaning validation. *PDA J Pharm Sci Technol PDA* 2000;**54**(5):365–72.
53. Bristol P. Cleaning Validation–The Rise of TOC. *Manuf Chem (England)* 2004;**75**:37–8.
54. Guazzaroni M, Yiin B, Yu JL. Application of total organic carbon analysis for cleaning validation in pharmaceutical manufacturing.. *Am Biotechnol Lab* 1998;**16**(10):66–7.
55. Biwald CE, Gavlick WK. Use of total organic carbon analysis and Fourier-transform infrared spectroscopy to determine residues of cleaning agents on surfaces. *J AOAC Int* 1997;**80**(5):1078–83.
56. Holmes AJ, Vanderwielen A. Total organic carbon method for aspirin cleaning validation. *PDA J Pharm Sci Technol* 1997;**51** (4):149–52.
57. Mirza T, Lunn MJ, Keeley FJ, George RC, Bodenmiller JR. Cleaning level acceptance criteria and a high pressure liquid chromatography procedure for the assay of Meclizine

Hydrochloride residue in swabs collected from pharmaceutical manufacturing equipment surfaces. *J Pharm Biomed Anal (England)* 1999;**19**:747–56.
58. Klinkenberg R, Streel B, Ceccato A. Development and validation of a liquid chromatographic method for the determination of amlodipine residues on manufacturing equipment surfaces. *J Pharm Biomed Anal* 2003;**32**(2):345–52.
59. Sajonz P, Natishan TK, Wu Y, McGachy NT, DeTora D. Development and Validation of a Sensitive and Robust Wipe-Test Method for the Detection and Quantification of the Antibiotic Ertapenem and its Primary Degradates in a Pharmaceutical Manufacturing Environment. *J Liq Chromatogr Related Technol* 2005;**28**(5):713–25.
60. Brindusa MB, Apostolides Z, Pretorius E. A validated HPLC method for determining residues of a dual active ingredient anti-malarial drug on manufacturing equipment surfaces. *J Pharm Biomed Anal* 2005;**37**(3):461–8.
61. Nozal MJ, Bernal JL, Toribio L, Jimenez JJ, Martin MT. Validation of the removal of acetylsalicylic acid. Recovery and determination of residues on various surfaces by high performance liquid chromatographic. *J Chromatogr A* 2000;**870**(1–2):69–75.
62. Nozal MJ, Bernal JL, Toribio L, Martin MT, Diez FJ. Development and validation of an LC assay for sumatriptan succinate residues on surfaces in the manufacture of pharmaceuticals. *J Pharm Biomed Anal* 2002;**30**(2):285–91.
63. Shea JA, Shamrock WF, Abboud CA, Woodeshick RW, Nguyen LQ, Rubino JT, et al. Validation of cleaning procedures for highly potent drugs. I. Losoxantrone. *Pharm Dev Technol* 1996;**1**(1):69–75.
64. Segretario J, Cook SC, Umbles CL, Walker JT, Woodeshick RW, Rubino JT, et al. Validation of cleaning procedures for highly potent drugs. II. Bisnafide. *Pharm Dev Technol* 1998;**3**(4):471–6.
65. Shifflet MJ, Shapiro M, Levin C, DeNisco R. Validating a reversed-phase HPLC method to determine residual nonoxynol-9 on pharmaceutical process equipment using a nonporous silica column. *LCGC North Am* 2001;**19**(3). 312, 314–317
66 Feketea S, Feketeb J, Ganzler K. Validated UPLC method for the fast and sensitive determination of steroid residues in support of cleaning validation in formulation area. *J Pharm Biomed Anal* 2009;**49**:833–8.
67. Kolodsick KJ, Phillips H, JFeng J, Molski M, Kingsmill CA. Enhancing drug development by applying LC-MS-MS for cleaning validation in manufacturing equipment. *Pharm Techno* 2006;**30**(2):56–71.
68. Liu L, Pack BW. Cleaning verification assays for highly potent compounds by high performance liquid chromatography mass spectrometry: strategy, validation, and long-term performance. *J Pharm Biomed Anal* 2007;**43**(4):1206–12.
69. Simmonds EL, Lough WJ, Gray M. Evaluation of LC-MS for the analysis of cleaning verification samples. *J Pharm Biomed Anal* 2006;**40**:631–8.
70. Payne K, Fawber W, Faria J, Buaron J, DeBono R, Azhar M, The Role of Spectroscopy in Process Analytical Technology, <www.spectroscopyonline.com>, 2005.
71. Debono R, Stefanou S, Davis M, Wallia G. Using ion mobility spectrometry for cleaning verification in pharmaceutical manufacturing. *Pharm Technol* 2002;**26**:72.
72. Munden R, Everitt R, Sandor R, Carroll J, Debono R. IMS limit test improves cleaning verification and method development. *Pharm Technol* 2002;**14**:66–72.
73 Strege M. Total residue analysis of swabs by ion mobility spectrometry. *Anal Chem* 2009;**81**(11):4576–80. Available from: http://dx.doi.org/10.1021/ac900441k.
74. Mirza T, George RC, Bodenmiller JR, Belanich SA. Capillary gas chromatographic assay of residual methenamine hippurate in equipment cleaning validation swabs. *J Pharm Biomed Anal* 1998;**16**(6):939–50.
75. Risley DS, Hostettler KF, Peterson JA. Trace analysis of a weak UV-absorbing pharmaceutical compound in swab samples using HPLC with evaporative light-scattering detection. *LC-GC* 1998;**16**(6):562–8.
76. Vovk I, Simonovska B. Development and validation of a thin-layer chromatographic method for determination of chloramphenicol residues on pharmaceutical equipment surfaces. *J AOAC Int* 2005;**88**(5):1555–61.
77. Katona Z, Vincze L, Vegh Z, Trompler A, Ferenczi-Fodor K. Cleaning validation procedure eased by using overpressured layer chromatography. *J Pharm Biomed Anal* 2000;**22**(2):349–53.
78. Fabre H, Altria KD. Key points in method validation in capillary electrophoresis. *LC-GC Europe* 2001;**14**(5). 302,304,306,308–310.
79. Altria KD, Creasey E, Howells JS. Routine capillary electrophoresis trace-level determinations of pharmaceutical and detergent residues on pharmaceutical manufacturing equipment. *J Liq Chromatogr Relat Technol* 1998;**21**:1093 30.
80 Altria KD, Hadgett TA. An evaluation of the use of capillary electrophoresis to monitor trace drug residues following the manufacture of pharmaceuticals. *Chromatographia* 1995;**40**:23.
81. Raghavan R, Mulligan JA. Low-level (PPB) determination of cisplatin in cleaning validation (rinse water) samples. I. An atomic absorption spectrophotometric method. *Drug Dev Ind Pharm* 2000;**26**(4):423–8.
82 Lewen N, Nugent D. The use of inductively coupled plasma-atomic emission spectroscopy (ICP-AES) in the determination of lithium in cleaning validation swabs. *J Pharm Biomed Anal* 2010;**52**:652–5.
83. Rotsch TD, Spanton M, Cugier P, Plasz AC. Determination of clarithromycin as a contaminant on surfaces by high-performance liquid chromatography using electrochemical detection. *Pharm Res (USA)* 1991;**8**:989–91.
84. Elrod Jr. L, Schmit JL, Morley JA. Determination of isoproterenol sulfate on surfaces using high-performance liquid chromatography with electrochemical detection. *J Chromatogr A* 1996;**723**:235–41.
85. Mehta NK, Goenaga-Polo J, Hernandez-Rivera SP, Hernandez D, Thomson MA, Melling PJ. *Spectroscopy (Duluth, MN, United States)* 2003;**18**(4). 14-16,18-19
86. Smith JM. *Pharm Technol (USA)* 1993;**17**. pp. 88, 90-92, 94, 96, 98
87. Urbas AA, Lodder RA. In Situ Spectroscopic Cleaning Validation. *NIR News* 2003;**14**(2):8–10.
88. Simonovska B, Andrensek S, Vovk I, Prosek M. High-performance thin-layer chromatography method for monitoring norfloxacin residues on pharmaceutical equipment surfaces. *J Chrom A* 1999;**862**:209–15.
89. Weston A. Ion chromatography in the pharmaceutical industry. *Am Biotechnol Lab* 1998;**16**(3). 30, 32–33.
90. Wilfredo R, Hernandez D, Rey R, Colon H, Zayas J. Cleaning validation 2: Development and validation of an ion chromatographic method for the detection of traces of CIP-100 detergent. *J Pharm Biomed Anal* 2007;**44**:265–9.
91. Shifflet MJ, Shapiro M. Validation of a Cation-Exchange Method to Test for Residual Amounts of a Cleaning and Sanitizing Solution on Pharmaceutical Process Equipment. *BioPharm* 2000;**13**(1). pp. 51-54 + 64.
92. Poster obtained from ESA. Forsatz, B.J., (Schering-Plough Research Institute), Snow, N.H. (Seton Hall University, Department of Chemistry and Biochemistry).
93. Pelesa DN, Elya KJ, Crowdera TM, Ponstinglb M. Rapid at-line pharmaceutical cleaning verification using a novel light induced fluorescence (LIF) sensor. *J Pharm Biomed Anal* 2013;**72**:1–7.
94 Zámostný P, Punčochová K, Vltavský Z, Patera J, Bělohlava Z. Dry-swabbing/image analysis technique for the pharmaceutical equipment cleaning validation. *Procedia Eng* 2012;**42**:447–53.
95. ICH Guidelines Q2 (R1). Validation of Analytical Procedures: Text and Methodology. Current Step 4 version Parent Guideline dated October 27, 1994.

CHAPTER

25

Specification Setting and Manufacturing Process Control for Solid Oral Drug Products

W. Chen[1], *S. Stithit*[2] *and J.Y. Zheng*[3]

[1]Gilead Sciences, Foster City, CA, United States [2]Century Pharmaceuticals, Indianapolis, IN, United States
[3]Ruby Pharma Inc., Duluth, GA, United States

25.1 INTRODUCTION

The quality of drug products is determined by their design, development, process controls, good manufacturing practice (GMP) controls, process validation, and by specifications applied to them throughout development and manufacturing.[1,2] As part of GMP controls, the quality of any drug product must be evaluated using methods and criteria listed in its specifications. According to ICH Q6A, a specification is defined as[1]:

> A list of tests, references to analytical procedures, and appropriate acceptance criteria that are numerical limits, ranges, or other criteria for the tests described. It establishes the set of criteria to which a drug substance or drug product should conform to be considered acceptable for its intended use. Conformance to specification means that the material, when tested according to the listed analytical procedures, will meet the listed acceptance criteria.

In the pharmaceutical industry, specifications are required by regulatory bodies for any drug substance, intermediates, and finished drug product. Specifications are critical to ensure product quality and consistency, and are an important part of an integrated quality strategy for finished drug products in addition to GMPs (eg, using suitable facilities, an established/validated manufacturing process, validated test procedures, raw materials testing, in-process testing, and stability testing).[1–3] Specifications for the drug substance, in-process tests, and final product release tests must be based on development information (eg, batch history, stability, efficacy, and toxicity) and certainly evolve with the cumulative knowledge gained during product research and development. In all cases, specifications either in the clinical trial stages or in the commercialization stage are chosen to confirm the quality of the drug product and, thus, must focus on the product characteristics that ensure the identity, strength, and purity of the drug product at release, as well as throughout the duration of clinical trial or the shelf-life for commercial products. It is important to have solid rationale and scientific justification for including and/or excluding testing for specific quality attributes because specifications are used to confirm the product quality rather than to characterize the product.

The following are four fundamental principles that should be taken into consideration during setting specifications for both drug substance and drug product:

1. Specifications should be based on relevant preclinical development and clinical study data. These clinical data define the boundaries of key safety and efficacy.
2. Specifications should be related to a manufacturing process. Specifications must be based on data obtained from batches used to demonstrate product performance and manufacturing consistency. Product specifications, especially for process-related degradation products, must be related to the specific manufacturing process.
3. The stability of drug substance and drug product should be taken into consideration for setting specifications. Physical and chemical degradation of drug substance and drug product may occur during storage. These changes could affect product safety and performance.

DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00025-X

4. Specifications must be linked to analytical procedures. Critical quality attributes (CQAs) may include characteristics such as assay, impurity profiles, and in vitro dissolution. Such attributes can be evaluated by multiple analytical methods to yield different results. During the course of product development, it is common for analytical technology to evolve in parallel with the product. Therefore, analytical data generated during development should be correlated with those generated at the time when the new drug application (NDA) is filed.

For the purpose of this document, early phase development refers to phase 1 or 2 clinical trials before the start of primary stability studies or pivotal clinical trials. Late phase development refers to phase 3 or pivotal clinical studies. Early phase and late phase specifications refer to specifications established during early phase and late phase development, respectively. The regulatory specification refers to specifications established for the registration and commercial batches.

The specification setting process begins with identification of a set of attributes for testing, development of analytical procedures, and establishment of an acceptance criteria that can be used to measure product quality. Specifications are then refined during product development and the commercialization process. Once specifications are set, process controls need to be identified in order to monitor, and/or adjust the manufacturing process. The goal of process control is to ensure that an in-process material or finished drug product meets its respective specifications.

Specifications and the specification setting process for drug substances, clinical trial materials, and drug products are described in this chapter. Process controls, particularly in-process material tests, statistical methodology for process controls, and application of process analytical technology (PAT) in-process controls are discussed. The impact of analytical procedures on the specifications is also discussed.

25.2 SPECIFICATIONS FOR THE DRUG SUBSTANCE

The drug substance, the active pharmaceutical ingredient (API), is the most important component in a drug product. During the early phase of development, the number of API batches is relatively limited, the batch size is usually very small, and manufacturing experience is beginning to evolve. The primary emphasis for setting specifications for a new drug substance is on safety based on the purposes of clinical studies at this stage. During the later phase of development as sufficient data and experience become available, product quality becomes the focus for setting specifications based on manufacturing processes and analytical capability, in addition to safety considerations.

ICH guideline Q6A generally applies to products during late stage development because the guideline specifically states that it is "not intended to apply to the regulation of new drug substance used during the clinical research stage of development."[1] Nonetheless, the main principles of the guideline should be considered carefully throughout development. As discussed previously, the primary focus of specifications is safety during the early phase of development. To set specifications at this stage, the tests described in ICH Q6A should be reviewed to determine which tests must have limits and which tests can be performed "to be monitored." To-be-monitored tests refer to those that do not have established limits or acceptance criteria. During the early phase of development, some tests can be performed "to be monitored" because whether or not the test is quality-indicating may not yet be known, or there may be insufficient data to set a proper limit. During the late phase of development, specifications are gradually developed toward meeting the requirements for registration. Therefore the late phase specifications should include tests expected at registration, although there may not be enough data available to set appropriate limits until just prior to submission.

Manufacturing process controls for drug substances are not discussed in this chapter. Please refer to the FDA guidance for industry ICH Q11 Development and Manufacture of Drug Substances issued in November 2012,[4] which is applicable to drug substances as defined in the scope sections of the ICH guidance documents Q6A and Q6B. ICH Q11 provides clarification on the principles and concepts described in ICH Guidelines on Pharmaceutical Development (Q8[R2]),[5] Quality Risk Management (Q9),[6] and Pharmaceutical Quality System (Q10)[7] as they pertain to the development and manufacture of drug substance.

The following tests and acceptance criteria are described in ICH Guideline Q6A. An example of early phase and late phase specifications are provided in Table 25.1.

- *Description*: A qualitative statement about the physical state and appearance of the new drug substance should be provided. For the example in Table 25.1, the acceptance criteria are based on the drug substance's historical visual appearance data.
- *Identification*: Identification testing should ideally be able to discriminate between compounds of closely related structure that are likely to be present. The test should be specific to the new drug substance

TABLE 25.1 Example of Specifications for Drug Substance

Tests	Method	Limits for early phase	Limits for late phase/commercial batches
Description	Visual	White to yellow solid	White to yellow solid
Identity by IR	IR	Spectrum compares favorably with that of the reference standard.	Spectrum compares favorably with that of the reference standard.
Identity by HPLC	HPLC	N/A	Retention Time compares favorably with that of the reference standard.
Assay	HPLC	Not less than 93.0% and not more than 105.0%	Not less than 93.0% and not more than 105.0%
Related substances	HPLC		
Total related substances		Not more than 3.0%	Not more than 2.0%
Individual related substances		Not more than 1.0%	Not more than 0.7%
Chiral impurities	HPLC		
Form A		Not more than 0.7%	Not more than 0.3%
Form B		Not more than 0.7%	Not more than 0.4%
Residual solvents	GC		
Solvent A		Not more than 2.0%	N/A
Solvent B		N/A	Not more than 0.5%
Water	Karl Fischer Titration	Not more than 5.0%	Not more than 2.0%
Residue on ignition	USP 38 (2015)	To be monitored	Not more than 0.05%
Heavy metals	USP 38 (2015)	To be monitored	Not more than 20 ppm
Particle size	Laser diffraction	To be monitored	N/A
Crystal form	X-ray diffraction	To be monitored	N/A
Microbiological test	Compendial method	N/A	Meet compendial requirements

such as infrared spectrum (IR). Identification solely by a single chromatographic retention time is not regarded as specific. However, the use of two chromatographic methods, where separation is based on different principles or a combination of tests into a single procedure, such as high-pressure liquid chromatography (HPLC)/UV (ultraviolet) diode array, HPLC/mass spectroscopy (MS), or gas chromatography (GC)/MS is generally acceptable. If the new drug substance is a salt, identification testing should be specific to the individual ions. An identification test that is specific for the salt itself should suffice. New drug substances that are optically active may also need specific identification testing or performance of a chiral assay. For an early phase specification, only one identification method is generally included. A spectroscopic method such as IR is generally used. For late phase specifications, two identification methods are typically included (see Table 25.1).

- *Assay*: This test is used to determine the purity of the drug substance. A specific, stability-indicating analytical method should be included to determine the content of the new drug substance. In many cases, it is possible to apply the same method (eg, HPLC) for both assay and quantitation of impurities. Both early and late phase specifications require a proper assay limit, which is commonly expressed as a percentage for new drug substances. The limit usually specifies a correction for water and/or solvent (ie, anhydrous, volatiles-free, or dried). For the example in Table 25.1, the assay method separates all known related substances and degradation products from the drug substance. The lower limit is calculated to allow for related substances and assay variability. The upper limit allows for assay variability. The specification for assay is the same for the drug substances at the early and late phases.

TABLE 25.2 Classification of Impurities in New Drug Substance

Categories	Physical state	Source	Examples
Organic impurities	Volatile; Nonvolatile	Manufacturing process; Storage	Starting materials By-products Intermediates Degradation products Reagents, ligands, and catalysts
Inorganic impurities	Nonvolatile	Manufacturing process	Reagents, ligands, and catalysts Heavy metals or other residual metals Inorganic salts Other materials (eg, filter aids, charcoal)
Residue solvents	Inorganic or organic liquid	Manufacturing process	Acetone (class 3) Acetonitrile (class 2) Benzene (class 1)

TABLE 25.3 Thresholds of Impurities in New Drug Substance

Maximum daily dose[a]	Reporting threshold[b,c]	Identification threshold[c]	Qualification threshold[c]
≤2 g/day	0.05%	0.10% or 1.0 mg/day intake (whichever is lower)	0.15% or 1.0 mg/day intake (whichever is lower)
>2 g/day	0.03%	0.05%	0.05%

[a]*The amount of drug substance administrated per day.*
[b]*Higher reporting thresholds should be scientifically justified.*
[c]*Lower thresholds can be appropriate if the impurity is unusually toxic.*

- *Impurities*: Categories of impurities in a new drug substance listed in Table 25.2 included organic and inorganic impurities, and residual solvents. Table 25.3 provides the identification and qualification threshold.[8]

  Organic impurities include starting materials, reaction by-products, process intermediates, degradation products, reagents, ligands, and catalysts. They are process-related except for degradation products, and may be identified or unidentified, volatile or nonvolatile. Chemical structures of impurities present in the new drug substance at or above an apparent level of 0.1% (eg, calculated using the response factor of the drug substance) should be characterized. The early phase specification of new drug substance for organic purities should include limits for total and any unspecified impurity (the largest reported unspecified impurity). Any identified impurity with a potential toxicological risk (eg, carcinogen, teratogen, mutagen) should also be included in the early phase specification with a limit that is toxicologically acceptable without risking human health. A qualitative comparison of impurity profiles, generally an HPLC impurities test, is also performed during the early phase. The impurity profiles of the batches of drug substance to be used in clinical trials must be compared with the batches used in toxicology studies and/or previous clinical trials. The late phase or regulatory specifications of the new drug substance for organic impurities should be consistent with the ICH Q3A guideline and include, where applicable, limits for:
  - Each specified identified impurity
  - Each specified unidentified impurity at or above 0.1%
  - Any unspecified impurity, with a limit of not more than 0.1%
  - Total impurities

  Table 25.1 shows the limits of not more than 1.0% of any individual related substance and not more than 3.0% of total related substances for the example drug substance during early phase development. The limits in this example are designed to ensure that the clinical trial material will not be significantly different, or, at a minimum, not inferior in terms of impurities from the material used in toxicological studies. The limits are tightened based on the process capability and stability data for the late phase specification.

  Inorganic impurities can result from the manufacturing process and usually are known and identified (eg, reagents, ligands and catalysts, heavy metals, inorganic salts, filter aids, and charcoal). These impurities are normally analyzed using pharmacopoeial or other appropriate methods. Limits could be based on

pharmacopoeial standards or known safety data. The example in Table 25.1 shows that residue on ignition and heavy metals were monitored without specifications established at the early stage of the investigation. The limits at late stage were based on historical data.

Solvents are inorganic or organic liquids used as vehicles for the preparation of solutions or suspensions in the synthesis of a new drug substance. For an early phase specification, a toxicology review of the manufacturing process may be needed to identify the solvents that require limits. In general, solvents that are used in the final purification step require an appropriate limit. Solvents that are particularly hazardous should be controlled prior to the final drug substance whenever possible. Solvents that likely to be, or already are, carried through to the final drug substance at a level of concern should be controlled carefully with a proper limit. At the late phase of development, any solvent that may exist in the drug substance should be quantified using appropriate analytical procedures. Control strategy must be in place if the solvents are generally known to be toxic per ICH guideline.[9,10] Control limits should be based on pharmacopoeial standards, ICH safety limits, or known safety data with consideration of clinical dose, duration of treatment, and route of administration. For example, in Table 25.1 the drug substance in early phase was recrystallized from solvent A. Material manufactured at the late phase was recrystallized from solvent B. The limits proposed for solvents A and B were based on the projected dose and the relative toxicity of these solvents.

ICH Q6A decision tree #1 shows how to establish an acceptance criterion for a specified impurity in a new drug substance.[1] Any limits to control impurities in the specification are based on the body of data generated during development. It is unlikely that sufficient data will be available to assess process consistency at the time of filing an NDA with the FDA. Therefore, acceptance criteria should not tightly encompass the batch data at the time of filing. Impurities present in the new drug substance and exceeding the ICH threshold listed in Table 25.3 should be qualified. Qualification is the process to evaluate in vivo toxicity of an individual impurity or a given impurity profile at a specified level. After successful testing in safety and/or clinical trials, the level of the impurity is considered qualified. Higher or lower threshold limits for qualification of impurities may sometimes be necessary for some drug categories. If an impurity of a drug or therapeutic class has been known to have adverse effects in humans, qualification threshold limits should be lower than usual. However, when the level of safety concern is less based on patient population, drug class effects, and clinical considerations, a higher threshold limit may be appropriate. The decision tree in attachment 3 of ICH Q3A discusses critical considerations for the qualification of impurities when threshold limits are exceeded. In general, information from literature should be first searched to qualify an impurity. In many cases, it can be easier to control impurities below the threshold limit during the manufacturing process than to conduct toxicological evaluation.

To set specification limits of impurities, all data from available stability studies, chemistry development, and scale-up studies, and batches for clinical trials should be used to predict potential impurities that may exist in the commercial product. Only impurities that are observed in batches manufactured by proposed commercial synthetic routes will be included in the specification of the new drug substance.

In addition to the universal tests described previously, the following tests may be considered on a case-by-case basis for new drug substances.[1]

- *Physicochemical properties*: These are properties such as pH of an aqueous solution, melting point/range, or refractive index.
- *Particle size and size distribution*: The ICH Q6A decision tree #3 provides general guidance on when particle size testing should be considered. Testing for particle size and size distribution should usually be performed if particle size can have a significant impact on dissolution rates, bioavailability, stability, and manufacturability. Table 25.1 shows that particle size was monitored, but a limit was not established for the drug substance at the early phase of the investigation. The test was removed at the late phase because particle size did not impact the dissolution rate of the drug substance and drug product, the stability of the drug substance, or manufacturability of the drug product. Therefore no specification needs to be established for particle size distribution.
- *Polymorphic forms*: Different polymorphic forms may affect drug product performance, bioavailability, or stability. Polymorphism includes different crystal forms of a new drug substance and its solvates, hydrates, and amorphous forms. Following the ICH Q6A decision #4(1) through 4(3), polymorphism of the new drug substance should be understood and controlled with appropriate physicochemical measurements and techniques. These parameters must be controlled until data show that control is not necessary (ie, only one form can be generated during polymorphs generation studies). Table 25.1 shows that the crystal form of the example drug substance was monitored, but a specification was not established at the early phase of the

investigation. The test was removed at the late phase because polymorphism has not been observed as a result of the current synthetic route and no specification needs to be established.

- Chiral drug substances present additional technical challenges to pharmaceutical scientists. ICH Q6A decision tree #5 summarizes when and whether chiral identity tests, impurity tests, and assays may be needed for both new drug substances and new drug products. For a drug substance developed as a single enantiomer, control of the other enantiomer should be considered in the same manner as other impurities. An enantioselective determination of the drug substance should be part of the specification. In addition, identity tests should be capable of distinguishing both enantiomers and the racemic mixture. It is important for a chiral drug substance to control starting materials or intermediates and enantiomeric impurity during the manufacturing process. The example in Table 25.1 shows that the limits at early phase are to ensure that lots are not significantly different from the quality of the lots for toxicology studies and allow for assay variability. The limits at late phase were tightened based on process capability and stability data available.
- *Water content*: Control of water content is especially important if the new drug substance is hygroscopic or degraded by moisture or when the drug substance is known to be stoichiometric hydrate.
- *Microbial limits*: At the early phase, microbial test limits may not be needed for new drug substances used for solid oral dosage forms. At the late phase of development, a microbial limit is generally required unless a scientific justification can be provided. Appropriate microbial tests and acceptance criteria are based on the nature of the drug substance, method of manufacture, and the intended use of the drug product. Tests may include the total count of aerobic microorganisms, the total count of yeasts and molds, and the absence of specific objectionable bacteria (eg, *Staphylococcus aureus*, *Escherichia coli*, *Salmonella*, *Pseudomonas aeruginosa*). ICH Q6A decision tree #6 provides additional guidance on when microbial limits should be included in the specifications of a drug substance and excipient.

25.3 SPECIFICATIONS FOR CLINICAL TRIAL MATERIALS

In each phase of a clinical investigational program, sufficient information is required to be submitted to the FDA to ensure the proper safety, identification, strength, quality, purity, and potency of the investigational candidate. However, the amount of information to provide varies with the phase and the proposed duration of investigation, and with the amount of information previously available. Specification for clinical trial materials is developed based on these factors: compendial standards, process development studies, and stability of batches used for toxicological/in vivo studies. Often the specification evolves as more manufacturing experience is gained.

25.3.1 Early development stage (Phases 1 and 2)

During early development, safety of the drug for use in clinical trials is the most important factor to consider in determining the specifications. When setting the specification limits for the drug substance, the toxicological lots are the benchmark for future lots. Specifications for impurities in early phase clinical trial materials can be based on the quality of materials used in toxicological studies and the known toxicities of impurities. In addition, intended clinical usage, available clinical data, analytical methods, and process variability should also be considered.

The FDA guideline for Phase 1 investigation[11–13] requires quantitative composition of the product, which includes a brief description of test methods to ensure the identity, strength, quality, purity and potency, test results, or a certificate of analysis. Specified quality attributes should be monitored based on applicable acceptance criteria. For known safety-related concerns, acceptance criteria should be established and met. For some product attributes, all relevant acceptance criteria may not be known at this stage. An impurity profile is recommended to be established to the extent that future reference and/or comparison is possible. However, not all product impurities need characterization at this stage.

During Phase 2 studies, physicochemical tests and microbiological tests that have been added or deleted from the specifications established for Phase 1 studies are required to be reported to the FDA.[14] Physicochemical tests include identity, assay, content uniformity, degradants, impurities, dissolution, and particle size. Data on the particle size distribution and/or polymorphic form of the drug substance used in clinical trial materials should be included so that correlations can be established between data generated during early and late phase drug development and in vivo performance. Relaxation of acceptance criteria or any changes that affect safety should also be reported.

25.3.2 Late development stage (Phase 3)

As development progresses into Phase 3, growing data on batch history, process capability, analytical

capability, and stability become increasingly important in the specification setting. An updated specification with a detailed listing of all the tests performed on the product and the tentative acceptance criteria should be provided to the FDA. Test results and analytical data from batch release of representative clinical trial materials should be provided initially and when any changes are made in the specification. Data on particle size distribution and/or polymorphic form of the drug substance should also be included. Degradation products should be identified, qualified, quantified, and reported. Suitable limits should be established based on manufacturing experience, stability data, and safety considerations. A dissolution testing program or drug release program should be developed for drug products.

Typical Phase 3 specifications for an oral drug product include description, identification, assay, content uniformity of dose units, degradation products, dissolution, water content, and microbiological tests. Additional in-process tests during clinical trial manufacturing of a tablet/capsule dosage form include: dimensions, disintegration, and weight (including hardness and friability for tablet dosage forms). The quality of the product used in these clinical studies will form the basis for product approval and the quality of the registered product has to reflect that of clinical trial materials.

An example of drug product (capsule) specifications at early phase and late phase development is provided in Table 25.4. The example specifications include description, identification, assay, related substances, water, dissolution, uniformity of dosage unit, and microbiological test. The limit for description is based on historical visual appearance data of the drug product. During early phase development, only an IR method is used for an identity test. In late phase development, two identification methods are used (ie, IR and HPLC methods). The upper/lower assay limit is calculated to allow for related substances and assay variability. The limit is typical and the same for early and late phase drug products. Related substances (total related substances and individual related substances) are "to be monitored" during the early phase of development. The limits are established for the late phase after more data became available. The USP limit is used for the uniformity of dose units test. Limit for dissolution is based on the dissolution study in previous historical batches. At the late phase of development, the limit is set at an earlier time point based on the data collected during early phase development.

TABLE 25.4 Example of Specifications for a Capsule Drug Product

Tests	Method	Limits for early phase	Limits for late phase
Description	Visual	White capsule	White capsule
Identity	HPLC	N/A	The retention time of the main peak in the chromatogram of the assay preparation corresponds to that in the chromatogram of the reference standard preparation, as obtained in the Assay testing
Identity	IR	The IR absorption spectrum of the sample preparation exhibits maxima only at the same wavelengths as that of corresponding reference standard	The IR absorption spectrum of the sample preparation exhibits maxima only at the same wavelengths as that of corresponding reference standard
Assay	HPLC	Not less than 90.0% and not more than 110.0% of label claim	Not less than 90.0% and not more than 110.0% of label claim
Related substances	HPLC		
Total related substances		To be monitored	Not more than 1.0%
Individual related substances		To be monitored	Not more than 0.5%
Water	Karl Fischer titration	To be monitored	5%
Uniformity of dosage units	HPLC	Meet USP 38 (2015)	Meet USP 38 (2015)
Dissolution	USP 38 (2015)	$Q = 75\%$ at 45 min	$Q = 75\%$ at 35 min
Microbiological tests	Compendial method	N/A	Meet compendial requirements

25.4 SPECIFICATIONS FOR COMMERCIAL DRUG PRODUCTS

Specifications for commercial drug products, also known as regulatory specifications, are generally documented in a regulatory submission or in a compendial monograph. Commercial drug product must conform to the regulatory specifications throughout its shelf life. Full compliance with the ICH guideline Q6A is expected for commercial drug products' specifications.

Specifications for commercial drug products are refined from late phase specification based on analytical capability, process capability, and product stability. These considerations augment the safety requirements. At the time of registration, the number of stability batches representative of the future commercial batches is often limited, and shelf-life stability testing has not been completed. Statistical extrapolation of limited stability data to the proposed shelf life is necessary. It is important that specifications are not set too tightly, which could result in unnecessary rejection and time-consuming investigations. On the other hand, controlling patients' risk of using products that do not meet requirements should be a top priority. Wessels et al. has proposed a strategy to statistically set an appropriate specification for pharmaceutical products.[15] The authors use linear regression to extrapolate the stability data available at the time of registration. They recommend that a confidence limit of 99% for individual results be used to determine specifications after batch-to-batch variability is tested according to the ICH guideline on stability testing.[16] Patient risk can be controlled by tightening the specification (eg, using a confidence limit of 95% instead of 99%).

The following universal tests described in ICH Q6A for drug product specifications are generally applicable to all new drug products:

- *Description*: A qualitative description of the finished product such as size, shape, and color should be provided.
- *Identification*: Identification testing is able to discriminate between compounds of closely related structures that might be present. Identity tests should be specific for the new drug substance. A single chromatographic retention time, for example, is not deemed specific. However, the use of two chromatographic procedures is acceptable if the separation is based on different principles. The combination of tests into a single procedure such as HPLC/UV diode array, HPLC/MS, or GC/MS is also acceptable.
- *Assay*: Assay methods should determine the drug product strength/content using a specific and stability-indicating method. In cases where use of nonspecific assay is justified, additional supporting analytical procedures should be used to attain overall specificity.
- *Impurities*: Impurities include organic and inorganic impurities, as well as residual solvents. Detailed information can be found in the following ICH guidelines: "Q3B (R2) Impurities in New Drug Products"[17] and "Q3C (R5) Impurities: Residual Solvents."[9] Organic impurities that are degradation products of the new drug substance and process-related impurities from the new drug product should be monitored. Acceptance limits for individual specified degradation products should be established. Table 25.5 provides the identification and qualification threshold. The emphasis is on

TABLE 25.5 Thresholds for Degradation Products in New Drug Products

Maximum daily dose[a]	Reporting threshold[b,c]	Maximum daily dose[a]	Identification threshold[b,c]	Maximum daily dose[a]	Qualification threshold[b,c]
≤1 g	0.1%	<1 mg	1.0% or 5 μg TDI, whichever is lower	<10 mg	1.0% or 50 μg TDI, whichever is lower
		1–10 mg	0.5% or 20 μg TDI, whichever is lower	10–100 mg	0.5% or 200 μg TDI, whichever is lower
>1 g	0.05%	>10 mg–2 g	0.2% or 2 mg TDI, whichever is lower	>100 mg–2 g	0.2% or 3 mg TDI, whichever is lower
		>2 g	0.10%	>2 g	0.15%

[a]*The amount of drug substance administrated per day.*
[b]*Thresholds for degradation products are expressed either as a percentage of the drug substance or as total daily intake (TDI) of the degradation product. Lower thresholds can be appropriate if the degradation product is unusually toxic.*
[c]*Higher reporting thresholds should be scientifically justified.*

impurities arising from drug product manufacture or from degradation. The impurity profiles of batches representative of the proposed commercial process should be compared with the profiles of batches used in development, and any differences should be discussed. The specifications of the new drug product for impurities should be consistent with the ICH Q3B guideline, which include, where applicable, limits for:

- Each specified identified degradation product
- Each specified unidentified degradation product
- Any unspecified degradation product, with a limit of not more than the identification threshold
- Total degradation products

The following additional tests described in the ICH Q6A are specific to solid oral drug products:

- *Dissolution*: Dissolution testing measures the rate of release of the drug substance from the drug product using a specific apparatus. Specification of dissolution should be set based on the data of bio-batches. Analysis of the quantity of drug released at a single time point is commonly used for immediate release (IR) formulations when the drug substance is highly soluble. On the other hand, appropriate test conditions and sampling procedures should be established for modified release formulations. Multiple time point sampling is appropriate for IR products of poorly soluble compounds and for extended or delayed release formulations (see ICH Q6A decision tree 7(1)). In cases where changes in dissolution rate have been demonstrated to significantly affect bioavailability, testing conditions that can discriminate between acceptable and unacceptable bioavailability are desirable (see ICH Q6A decision tree 7(2)). In vitro–in vivo correlation may be used to establish acceptance criteria for extended release drug products when human bioavailability data are available for formulations exhibiting different release rates. Acceptance criteria should be established based on available batch data when bioavailability data are not available and drug release is not shown independently of in vitro test conditions. The variability in mean release rate at any given time point should not exceed a total numerical difference of $\pm 10\%$ of the labeled content of the drug substance unless a wider range is supported by a bioequivalency study (see ICH Q6A decision Tree 7(3)).
- *Disintegration*: Disintegration testing can be substituted for dissolution for rapidly dissolving products containing drugs that are highly soluble throughout the physiological pH range. Disintegration is most appropriate when a relationship to dissolution has been demonstrated or when disintegration is shown to be more discriminating than dissolution (see ICH Q6A decision tree 7(1)).
- *Hardness/friability*: Hardness/friability testing is usually performed as an in-process control. The test should be included in the specification only when the characteristics of hardness/friability have a critical impact on product quality (eg, chewable tablets).
- *Uniformity of dosage unit*: Testing the uniformity of dosage units includes measurement of both the mass of the dosage form and the content of the active ingredient in the drug product. Pharmacopoeial methods should be used. Often this test is performed as an in-process evaluation. In Dec. 2011, USP published an alternative condition that allows hard capsules, uncoated tablets, and film-coated tablets not meeting the 25 mg/25% requirement to be tested by the weight variation procedure if the relative standard deviation of the drug substance in the final dosage unit is not more than 2% demonstrated in process validation or development data. In Jun. 2014, the FDA issued Guidance for Industry, Q4B Evaluation and Recommendation of Pharmacopoeial Texts for Use in the ICH Regions Annex 6 Uniformity of Dosage Units General Chapter.
- *Water content*: A quantitation procedure that is specific for water is preferred (eg, Karl Fischer titration). A loss on drying procedure may be adequate in some cases. Acceptance criteria should be based on data regarding the effects of hydration or water absorption on the drug product. Table 25.6 shows a water specification of not more than 7% for an example tablet product. The water in this product is mainly contributed by the water of hydration from excipients. Primary stability data indicate that the packaged tablets absorb water under the long-term storage condition. Based on the stability data, the absolute change through the expiration-dating period and uncertainty in the estimated change are statistically computed. The water specification of not more than 7% is then established based on the process capability, change, and uncertainty of the change estimated from the stability data, as well as the method variation.
- *Microbial limits*: The type of microbial testing and acceptance criteria used should be based on the nature of the drug, method of manufacture, and the intended use of the drug product. Such testing may not be necessary with acceptable scientific justification (see ICH Q6A decision tree 8).

TABLE 25.6 Example of In-House Release and Regulatory Specification for a Tablet Product

Tests	Method	In-house release limits	Regulatory limits
Description	Visual	Colored, shaped, and debossed tablet	Colored, shaped, and debossed tablet
Identity by IR	IR	The IR absorption spectrum of the sample preparation exhibits maxima only at the same wavelengths as that of corresponding reference standard	The IR absorption spectrum of the sample preparation exhibits maxima only at the same wavelengths as that of corresponding reference standard
Identity by HPLC	HPLC	The retention time of the main peak in the chromatogram of the assay preparation corresponds to that in the chromatogram of the reference standard preparation, as obtained in the Assay testing	The retention time of the main peak in the chromatogram of the assay preparation corresponds to that in the chromatogram of the reference standard preparation, as obtained in the Assay testing
Assay	HPLC	Not less than 97.0% and not more than 106.5% of label claim	Not less than 90.0% and not more than 110.0% of label claim
Related substances	HPLC		
Total related substances		Not more than 1.09%	Not more than 2.0%
Individual related substances		Not more than 0.52%	Not more than 1.0%
Water	Karl Fischer titration	Not more than 3.8%	Not more than 7.0%
Uniformity of dose units	HPLC	Meet USP 38 (2015)	Meet USP 38 (2015)
Dissolution	USP 38 (2015)	$Q = 75\%$ at 30 min	$Q = 75\%$ at 30 min

25.4.1 Product in-house release specifications and regulatory specifications

Regulatory specifications are legal commitments that a commercial product must conform to throughout its shelf life. In order to meet the regulatory specification, more restrictive in-house release specifications are commonly established for drug products. For examples, tighter assay and impurity (degradation product) in-house release limits can be established to ensure that a product can meet regulatory specifications at the end of its shelf life. In Japan and the United States, the concept of release specification may only be applicable to in-house criteria, and not to the regulatory release criteria. In these regions, the regulatory acceptance criteria are the same from release throughout shelf life. However, many companies choose to have tighter in-house limits at the time of release to provide assurance that the product will remain within the regulatory acceptance criterion throughout its shelf life. In the European Union, distinct specifications for release and for shelf life are a regulatory requirement.

Degradation of drug product needs to be factored into specification setting, particularly for analytical properties (eg, potency and impurities) that are affected by the degradation. The in-house release limits for quantitative analytical properties can be calculated from regulatory limits based on the stability change (increase or decrease) and the measurement variability. The calculated in-house release limits are narrower than regulatory limits to compensate for measurement variability and/or product instability. The buffer between regulatory limit and in-house limit typically consists of the estimated absolute change through the expiration-dating period (change), uncertainty in estimated change (S_{change}), and measurement variation (S_{m}).

For an analytical property that degrades during storage, the upper and lower in-house limits (UL and LL) can be calculated from upper and lower regulatory limits (URL and LRL) using the following equations. Allen et al. used a similar equation to calculate the setting of in-house limits for a product that degrades during stability storage.[18]

$$UL = URL - t \times \sqrt{S_{\text{m}}^2}$$
$$LL = LRL + \text{Change} + t \times \sqrt{S_{\text{change}}^2 + S_{\text{m}}^2}$$

Where, t is determined from a t-table using appropriate degree of freedom and level of confidence.

For an analytical property that increases during stability storage, the UL and LL can be calculated from URL and LRL using these equations:

$$UL = URL - \text{Change} - t \times \sqrt{S_{\text{change}}^2 + S_{\text{m}}^2}$$
$$LL = LRL + t \times \sqrt{S_{\text{m}}^2}$$

For an analytical property absent of instability, the buffer between in-house limits and regulatory limits typically consists of measurement variability only.

An example of in-house release and regulatory specification for a tablet product is provided in Table 25.6. The assay limits in Table 25.6 are based on the safety, efficacy, process variability, and assay variability. The upper/lower in-house release limits for assay are calculated to protect against exceeding the regulatory limits from stability and assay variability. The regulatory limits of not more than 1.0% of any individual related substance and not more than 2.0% of total related substances are the same for drug substance because there is not significant degradation observed during drug product manufacturing process. The UL/LL are calculated based on stability and assay variability to protect the regulatory limits.

25.4.2 Product stability and expiration date

Once specifications are set, the expiration-dating period of a drug product is then established based on the stability data obtained from long-term stability studies. The FDA guideline[16] requires that at least three batches be tested to allow for a good estimate of batch-to-batch variability. Stability data from all batches should be combined to estimate an expiration-dating period for a drug product. Batch similarity should be tested before all stability data are combined for statistical analysis. Batch similarity is examined by testing the similarity of intercepts and slopes of degradation lines from different batches. If the hypotheses of comparable intercepts and slopes are not rejected at the 0.25 level of significance, the expiration-dating period can be estimated based on the pooled stability data of all batches. On the other hand, if the hypotheses of comparable intercepts and slopes are rejected at the significance of 0.25, the expiration-dating period should be estimated based on the shortest expiration-dating period calculated from the worst batch.

An estimate of an expiration-dating period is illustrated in the following example: four different batches (ie, Lots A, B, C, and D) of a single tablet formulation were placed on long-term stability. Lot B completed 36-month stability. Lots A, C, and D completed 24-month stability. The analysis assumes that zero-order change (linear modeling on the original data scale) is appropriate. Before pooling the data from the four lots to estimate the shelf life, a preliminary statistical test was performed to determine whether the regression lines from different lots have a common slope and/or time zero intercept. The statistical model with the effects of age, lots, and age × lots was used. The probability associated with the interaction term age × lots was 0.0063, which is smaller than 0.25. This indicates that there is sufficient evidence that the slopes across lots are statistically different and the prediction for each lot must be evaluated separately. The individual slope and 95% confidence intervals for each lot are then estimated. The values for both the upper and lower confidence intervals were compared with the regulatory specifications of 95−105%. The confidence interval with the shortest expiry dating was identified as Lot A. The projected expiration-dating period of Lot A supports the 36-month dating. Therefore the expiration-dating period of the product is set at 36 months.

For more details on shelf life determination, readers should refer to chapter "Statistical Design and Analysis of Long-Term Stability Studies for Drug Products."

25.5 PROCESS CONTROL FOR SOLID ORAL DRUG PRODUCTS

Process control is an all-inclusive term used to describe the controls used during production to monitor and, if appropriate, adjust the process, and/or to ensure that an in-process material with an established specification or the finished drug product will conform to its respective specification.[2] By using quality by design principles during process and product development, a robust process can be developed. Per ICH Q8(R2) and Q9, the relationship between critical process parameters and CQAs are obtained, and the acceptable operating ranges are established during process development. Using this established operating range of process parameters; the product meeting required quality attributes can be produced. As stated in ICH Q8(R2), working within this operating range of process parameters is not considered a change because it still produces the product meeting required quality attributes. This chapter will not discuss how to establish the ranges for process parameters. Four types of process controls are:

- *Operating parameters*: Conditions that can be adjusted to control the manufacturing process (eg, temperature, pH, mixing time, and mixing speed)
- *Environmental controls*: Conditions associated with the manufacturing facility (eg, facility temperature, humidity, clean room classification)
- *Process tests*: Measures used to monitor and assess the performance of the process (eg, product temperature or exhaust temperature during drying in fluid-bed drier)
- *In-process material tests*: Measures used to assess the quality attributes of an in-process material and ultimately decide to accept or reject the in-process material or drug product

Steps in the manufacturing process should have the appropriate process controls identified. All in-process material tests and any of the operating parameters,

environmental conditions, and process tests that ensure each critical manufacturing step is properly controlled should be established in order to meet the predetermined specifications of the final drug product. All critical process controls and their associated ranges, limits, or acceptance criteria should be identified and justified by experimental data. For critical operating parameters and environmental controls, numeric ranges, limits, or acceptance criteria typically can be based on the experience gained during the development of the manufacturing process.

25.5.1 In-process material tests and quality attributes

In-process material tests are one of the critical components of the process controls. They may be performed during the manufacture of drug substance or drug product, rather than as part of the formal battery of tests that are conducted prior to product release.[1] In-process tests that are only used for the purpose of adjusting process parameters within an operating range (eg, hardness and friability of tablet cores that will be coated and individual tablet weight) are not included in the specifications. Table 25.7 lists typical in-process material tests for solid oral dosage forms.

The process material tests directly assess the quality attributes of the in-process materials. Quality attributes are important chemical, physical, and microbiological properties or characteristics of the drug product and intermediate material that must be maintained for proper product performance. Quality attributes of the process are those process parameters that are controlled to ensure the process reproducibly meets the drug product quality attributes. These attributes are identified during product and process development and become the basis for ensuring the control of product quality.

During the compression/encapsulation process, in-process material testing is performed on solid oral dosage forms to assure consistency throughout these unit operations. Moisture content is typically performed for granules before tableting and encapsulation. Tablet weights, hardness, thickness, friability, and disintegration testing are typically performed for tablets. Weight testing is performed for capsules. Hardness and disintegration specifications are established during development and bio-batch production. Testing performed during commercial production is to demonstrate both comparability and consistency.

All in-process material tests are critical process controls because they directly assess the quality attributes of an in-process material and ultimately lead to a decision to accept or reject the in-process material or drug product. Well-defined proven acceptable ranges (PARs) should be established for these tests, so that the impact on product quality of operating outside the PARs is understood or can be predicted. Data from in-process tests should be routinely evaluated. Statistical tools such as control charts can be used to aid the evaluation process.

TABLE 25.7 Examples of In-Process Tests for Solid Oral Dosage Forms

Process step	In-process tests	Effects on process and quality attributes
Drying in wet granulation	Loss on drying	• Product stability
Blending/mixing	Blending uniformity	Uniformity of dose unit
Blending/mixing	Particle size distribution	Uniformity of dose unit
Blending/mixing	Bulk and tapped density[a]	Uniformity of dose unit
Blending/mixing	Flow ability	Uniformity of dose unit
Tablet compression	Weight	Uniformity of dose unit
Tablet compression	Thickness	Disintegration/dissolution
Tablet compression	Hardness	• Coating quality • Disintegration/dissolution
Tablet compression	Friability[b]	• Coating quality • Potency • Shipping
Capsule filling	Weight	Uniformity of dose units
Capsule filling	Joint length	Shipping (breakage due to improper lock) Packaging (blistering)
Tablet/capsule	Disintegration	Dissolution

[a]*FDA. Guidance for Industry Q4B Evaluation and Recommendation of Pharmacopoeial Texts for Use in the ICH Regions Annex 13 Bulk Density and Tapped Density of Powders General Chapter; May 2013.*
[b]*FDA. Guidance for Industry Q4B Evaluation and Recommendation of Pharmacopoeial Texts for Use in the ICH Regions Annex 9 Tablet Friability General Chapter; April 2010.*

25.5.2 Powder blending uniformity

Powder blending is a fundamental process of mixing for solid dosage forms to ensure the content uniformity of final drug products.[19–21] Current GMPs in 21CFR PARTS 210 and 211 state "such control procedures should include adequacy of mixing to assure uniformity and homogeneity." In 2003, the FDA issued a guidance document on powder blends and finished dosage units to promote a science-based policy and regulatory enforcement.[22,23] The guidance

described the procedure for assessing powder-mix adequacy, correlating in-process dosage unit test results with powder blend test results, and establishing the initial criteria for control procedures used in routine product manufacturing. However, in Aug. 2013, the FDA withdrew the draft guidance document for industry, Powder Blends and Finished Dosage Units—Stratified In-Process Dosage Unit Sampling and Assessment. The guidance document was withdrawn because sections V and VII of the guidance document no longer represented the FDAs current opinion.[24] Section V (Exhibit/Validation Batch Powder Mix Homogeneity) recommended that at least 3 replicate samples be taken from at least 10 different locations in the powder blender, but only 1 of the 3 replicates be evaluated to assess powder blend uniformity. The FDA currently prefers that all replicate samples taken from various locations in the blender be evaluated in order to demonstrate the homogeneity of powder blend by using a statistically valid analysis. The acceptance criteria for Section VII (Routine Manufacturing Batch Testing Methods) were based on the limits published in the United States Pharmacopeia (USP) General Chapter <905> Uniformity of Dosage Units. However, the results should not be extrapolated to a larger population since USP <905> does not use a statistical sampling plan. The procedure and acceptance criteria in Section VII provided only limited statistical assurance that batches of drug products met appropriate specifications and statistical quality control criteria. Therefore, the FDA does not support the use of the results for batch release. The withdrawal of the document resulted in some uncertainty for manufacturers; however, manufacturers can still demonstrate blend uniformity as required per GMP using science and risk-based approach. The FDA recommends that ASTM E2709,[25] "Standard Practice for Demonstrating Capability to Comply with as Acceptance Procedure" and ASTM E2810,[26] "Standard Practice for Demonstrating Capability to comply with the Test for Uniformity of Dosage Units" should be used for application such as blend uniformity and uniformity of dosage unit. International Society for Pharmaceutical Engineering (ISPE) recently sponsored Blend Uniformity and Content Uniformity Working Group published recommendations for the assessment of blend and content uniformity with modifications to withdrawn FDA draft stratified sampling guidance.[27] The ISPE article focuses on statistical analyses assuming traditional sampling and analytical techniques that are widely used in the pharmaceutical industry for the assessment of blend and dosage unit uniformity. It recommends that statistical approaches and acceptance criteria including the level of confidence and coverage associated with the method should be selected using a risk-based approach that balances consumer's and manufacturer's risks, and other factors such as the therapeutic index of the drug. Two statistical approaches that may be used are the ASTM E2709/E2810 method and tolerance interval approach. The paper also links testing of blend and content uniformity with three stages of validation throughout product life cycle, including process design, process qualification, and continued process verification stages.

Powder blend uniformity is a process control (ie, in-process material test) that critically impacts the uniformity of dosage form. In-process sampling and testing is essential to determine blend uniformity. The sampling method and technique is crucial to obtain samples that adequately represent the powder mix. The traditional sampling device is a powder sample thief. However, certain sampling techniques may falsely introduce significant variations. The orientation, angle, and depth of sampling thief insertion, as well as insertion force and smoothness, impact the consistency of the sampling. Potential powder segregation during powder transfer and storage, and powder properties (eg, flowability, particle size distribution, and density) also add to the challenge of obtaining representative in-process samples. Therefore, the reliability of the sampling method should be evaluated as part of analytical methods development if a thief sampling method is used. The FDA encourages pharmaceutical companies to adopt a science and risk-based sampling approach, as well as more innovative technology such as PAT and statistical process control to ensure powder blend uniformity. To address sampling variations, the FDA recommends the use of a stratified sampling scheme when the batch contains locations that may have different results for a measured quality characteristic. Stratified sampling is the process of collecting a representative sample by selecting units deliberately from various identified locations within a lot or batch, or from various phases or periods of a process. With stratified sampling, a sample dosage unit is obtained to specifically target locations throughout the compression/filling operation that have a higher risk of producing failing results in the finished product tests for uniformity of content. This sampling method allows for collection and analysis of powder samples and dosage unit samples. Between- and within-location variability in the powder blend is a crucial component of finished product homogeneity and quality. The sampling and testing should demonstrate that no differences exist between locations in a powder blend that could adversely affect finished product quality. The sample variability can be attributed to either the blend's lack of uniformity or sampling error. Significant within-location variance in the

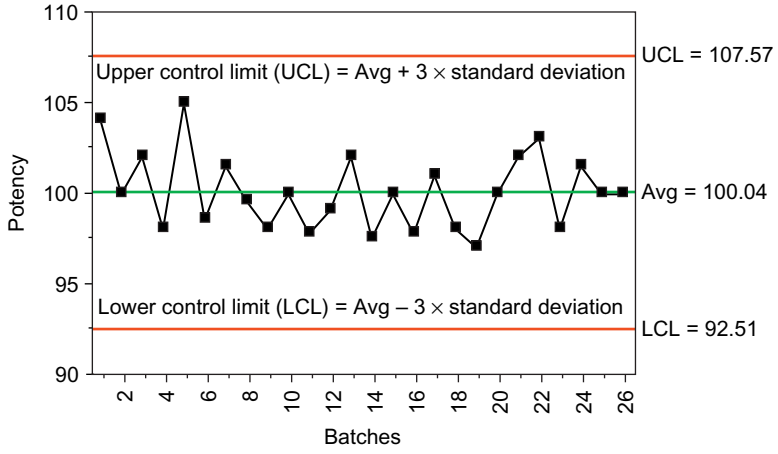


FIGURE 25.1 Example of control chart.

blend data represents sampling error, while high between-location variance in the blend data and high variance over time of one location can indicate that the blending operation is either inadequate, or sampling is biased, or the combination of inadequate mixing and sampling bias.

25.5.3 Statistical methodology for process control

In-process material testing results can be graphed in a control chart for routine evaluation. A control chart is a line graph displaying the results of repeated sampling of a process over time together with the control limits statistically established from historical process data. Control charts are commonly used to monitor the quality attributes and process stability. Control charts are one of the most important and effective statistical tools for quality control to monitor the process performance and trend detection. They can be applied for both in-process testing and finished product testing.

Process variability is the result of common cause variation and special cause variation. Common cause variation is a normal and consistent variation that occurs regularly. It is the random fluctuation expected in testing and follows a predictable distribution over time. The following are some examples of common cause variations:

- Environmental (eg, humidity, temperature) changes
- Baseline noise caused by regular electric interference
- Normal drift in a balance

Special cause variation is inconsistent and abnormal variation. It happens infrequently, and is an unexpected event. The only way to reduce special-cause variation is to investigate its cause and then implement countermeasures to prevent the cause from reoccurrence. The following are some of examples of special cause variations:

- Power outages causing changes in stability chamber storage conditions
- Equipment failure
- Incorrectly prepared stock standard in analysis

An example of a control chart is provided in Fig. 25.1. The centerline of the graph indicates the mean. The control limits typically include mean (μ) and $3\times$ standard deviation (σ) lines. The control chart limits represent the common variation of a process and are established from historical data. If process data are approximately normal, approximately 99.73% of the data will be located within ± 3 standard deviation from the process mean. Values falling outside of these limits are alarming signals because statistically individual values from a normal distribution will be within 3.0 standard deviations of the mean 99.73% of the time. If a process is only subject to common cause variation, a value more than three times the standard deviation is highly unlikely (only 2.7 occurrences per 1000 observations). Therefore, a value outside of control chart limit can be considered an indication of special cause variation. An example of control chart with a value outside of limit is provided in Fig. 25.2.

In addition to a value outside of control chart limits, special cause variation is shown by one or more of the following common out-of-control alarms[28]:

- 7 sequential points are on one side of the mean (a run)
- 7 sequential points increase or decrease steadily (a trend)

A run of 7 is probably not random statistically, assuming the observed values are from a normal distribution. The probability of 7 consecutive values being below the mean is $(1/2)^7 = (1/128) = 0.8\%$. The occurrence of such an event often suggests that the process

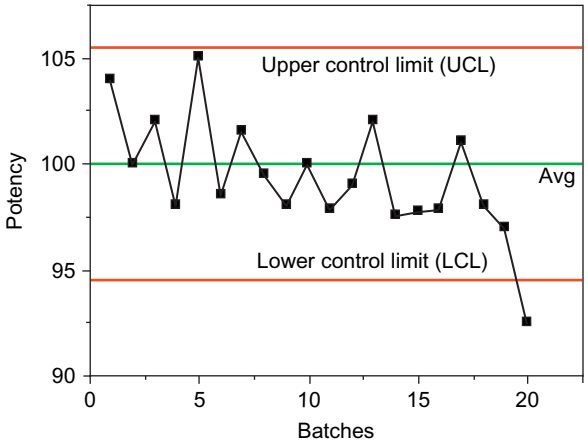


FIGURE 25.2 Example of control chart with a value outside of the limit.

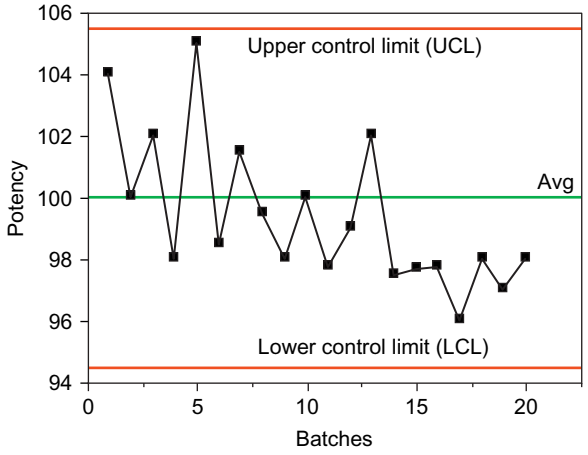


FIGURE 25.3 Example of control chart with 7 or more sequential points on one side of the mean.

should be carefully investigated. An example of control chart with 7 or more sequential points on one side of the mean is provided in Fig. 25.3.

The goal of control charting is to separate variability into common-cause and special-cause, with the knowledge that special-cause variability can be addressed and eliminated.

When a control chart alarm is observed, an investigation should be performed to identify the assignable cause.

25.5.4 PAT and in-process controls

The FDA's PAT Guidance[29] and 21st Century GMP Initiatives[30] promote a technology-based regulatory environment that encourages the pharmaceutical industry to implement innovative development, manufacturing, and quality assurance practices. The agency is encouraging the industry to utilize science and engineering-based principles to characterize and understand their processes. With this understanding, a manufacturing process is expected to become more efficient, consistent, and robust, providing the scientific rationales for the development of product specifications. For characterization and control purposes, PAT is a promising technology platform to accomplish these goals.

A PAT platform consists of an automated analyzer system that is integrated with process equipment for operation in a manufacturing environment. The system is capable of generating, analyzing, displaying, and transmitting product and process trend data continuously as a process unit operation is executed. For example, PAT can be used to monitor the critical unit operations of drying and blending to ensure optimal product dryness and the uniformity during blending to allow a homogenous blend for forward processing. This results in a more consistent product quality, faster forward processing and batch release. Wildfong et al. demonstrated that near-infrared (NIR) can be used to monitor the moisture content of the granules as they dry in real time.[31] Frake et al. showed that continuous in-line NIR monitoring of water content and particle size data are suitable means of process control and end-point determination for a fluid-bed granulation drying process.[32] Hailey et al. also demonstrated that NIR spectroscopy interfaced with process equipment and in combination with a fully automated system, can be used for real-time on-line powder blend homogeneity analysis.[33]

PAT can also be used in finished product analysis. Cogdill et al. developed and validated NIR models for the analysis of API content and hardness of tablets.[34] Spencer et al. demonstrated the feasibility of predicting active drug content in commercial-grade pharmaceutical tablets by NIR spectroscopy.[35] A PAT platform of at-line or in-line tablet analysis provides early warning in case of problems. It also offers a near real-time confirmation that the finished product is in compliance with product specifications. More detailed discussions of PAT and real-time release testing applications in the development and commercial manufacturing of oral dosage forms are provided in chapter "Commercial Manufacturing and Product Quality" of this book.

25.6 ANALYTICAL PROCEDURES

Analytical procedures evolve with the course of the product development. Analytical procedures may also change as knowledge of process controls (eg, what the critical impurities are) increases. Comparability of

results from changed analytical procedures must be considered during specification setting. It is important to confirm that data generated during development correlate with those generated at the time NDA is submitted.

An appropriate performance of the analytical procedures is critical for specification setting. For potency assay, the variability of the procedure must be taken into account for the specification setting in addition to the primary concern of patient safety and batch-to-batch variability. An intermediate precision study may be used as an estimate for the assay variability. For impurity and residual solvent testing, the requirements are outlined in the ICH guidelines.[8,9,17] The testing procedure must be able to determine the levels of impurities with appropriate accuracy and precision. For a specification limit of 0.1% for unknown impurities, the ICH reporting threshold of 0.05% can be used as the required quantitation limit to provide a reliable quantitation at the specification limit. For impurities with potential toxicity, toxicology should be consulted early in the drug development process to ensure methods with appropriate detection limits are developed and validated.

Analytical procedures must be validated for setting specifications. The performance and validation of a particular analytical procedure is influenced by other testing procedures used in setting the specification because a drug substance or drug product is tested by a set of complementary and supplementary testing procedures. It may not be necessary to optimize an individual analytical procedure to its best performance as long as it, together with other testing procedures, adequately tests the drug substance or product.

Validation of analytical procedures is an important part of the registration application for a new drug. chapter "Analytical Development and Validation for Solid Oral Dosage Forms" of this book discusses development and validation of analytical procedures in more detail.

25.7 CONCLUSIONS

Specifications are based on relevant development data, pharmacopoeial standards, test data of drug substances and drug products used in toxicology and clinical studies, and results from accelerated and long-term stability studies. In addition, analytical and manufacturing variability should be taken into consideration. Since specifications evolve through development and commercialization phases, specifications should be evaluated and updated periodically as additional data become available. Rationale for changes to specifications must be scientifically sound, and documented to allow traceability to previous versions.

To ensure specifications are met and product quality attributes are maintained, process controls must be identified, implemented, and effectively controlled during the manufacturing process.

Setting specifications is to establish the quality criteria for drug substance and product. The quality of drug product cannot be tested into the product. Instead quality should be built in by design. Setting proper specifications and maintaining effective process controls requires mechanistic understanding of how formulation and process design impact product quality and performance. Using science-based and risk-based approaches, the knowledge gained during different phases of development can be used to justify the tests and criteria in specifications to ensure the safety, efficacy, and quality of drug substances and products.

Acknowledgments

The authors would like to acknowledge the contributions of Richard Hwang of Pfizer to the previous version of this chapter.

References

1. FDA. International Conference on Harmonization; Guidance on Q6A specifications: test procedures and acceptance criteria for new drug substances and new drug products: chemical substances. *Fed Regist* 2000;**65**(251):83041–63.
2. FDA. Guidance for Industry INDs for Phase 2 and Phase 3 Studies Chemistry, Manufacturing, and Controls Information. Available from: www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/; May 2003.
3. FDA. Guidance for Industry Q7A Good Manufacturing Practice Guidance for Active Pharmaceutical Ingredients. Available from: www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/; August 2001.
4. FDA. Guidance for Industry, Q11 Development and Manufacture of Drug Substances, ICH; November 2012.
5. FDA. Guidance for Industry, Q8(R2) Pharmaceutical Development, ICH; November 2009.
6. FDA. Guidance for Industry, Q9 Quality Risk Management, ICH; June 2006.
7. FDA. Guidance for Industry, Q10 Pharmaceutical Quality System, ICH; April 2009.
8. FDA. Guidance for Industry Q3A Impurities in New Drug Substances, ICH. Available from: www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/; June 2008.
9. FDA. Guidance for Industry Q3C-Tables and List ICH; February 2012. FDA. Guidance for Industry Q3C Impurities: Residual Solvent. ICH. Available from: www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/; December 1997. ICH Q3C (R5) Impurities: Residual Solvent, ICH; February 2011.
10. FDA. Guidance for Industry Q6B Specifications: test procedures and acceptance criteria for biotechnological/biological products. Available from: www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/; August 1999.
11. Guidance for Industry CGMP for Phase 1 Investigational Drugs. Available from: www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/; July 2008.

12. Guidance for Industry, Investigators, and Reviewers Exploratory IND Studies. Available from www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/; January 2006.
13. Guidance for Industry Content and Format of Investigational New Drug Applications (INDs) for Phase1 Studies of Drugs, Including Well-Characterized, Therapeutic, Biotechnology-derived Products. Available from: www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/; November 1995.
14. Guidance for Industry INDs for Phase 2 and Phase 3 Studies Chemistry, Manufacturing, and Controls Information. Available from: www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/; May 2003.
15. Wessels P, Holz M, Erni F, Krummen K, Ogorka J. Statistical evaluation of stability data of pharmaceutical products for specification setting. *Drug Dev Ind Pharm* 1997;**23**(5):427–39.
16. FDA. Guidance for Industry Q1A (R2) Stability Testing of New Drug Substances and Products, ICH. Available from: www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/; June 2006.
17. FDA. Guidance for Industry Q3B (R2) Impurities in New Drug Products, ICH. Available from: www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/; July 2006.
18. Allen PV, Dukes GR, Gerger ME. Determination of release limits: a general methodology. *Pharm Res* 1991;**8**:1210.
19. Berman J. The compliance and science of blend uniformity analysis. *PDA J Pharm Sci Technol* 2001;**55**(4):209–22.
20. Venables HJ, Wells JI. Powder mixing. *Drug Dev Ind Pharm* 2001;**27**(7):599–612.
21. Geoffroy JM, LeBlond D, Poska R, Brinker D, Hsu A. Unit dose sampling and final product performance: an alternative approach. *Drug Dev Ind Pharm* 2001;**27**(7):731–43.
22. FDA. Guidance for Industry: Powder Blends and Finished Dosage Units—Stratified In-Process Dosage Unit Sampling and Assessment; October 2003. FDA Withdrawal Guidance; August 2013.
23. Boehm G, Clark J, Dietrick J, Foust L, Garcia T, Gavini M, et al. The use of stratified sampling of blend and dosage units to demonstrate adequacy of mix for powder blends. *PDA J Pharm Sci Technol* 2003;**57**:59–74.
24. FDA. Questions and Answers on Current Good Manufacturing Practices, Good Guidance Practices, Level 2 Guidance—Production and Process Controls. Available from: www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/.
25. ASTM E2709, Standard Practice for Demonstrating Capability to Comply with as Acceptance Procedure. Current edition approved October 1, 2014, editorial updated March 2015.
26. ASTM E2810, Standard Practice for Demonstrating Capability to Comply with the Test for Unifoirmity of Dosage Units. Current edition approved October 1, 2011, editorial updated December 2013.
27. Garcia T, Bergum J, Prescott J, Tejwani R, Parks T, Clark J, et al. Recommendations for the assessment of blend and content uniformity: modifications to withdrawn FDA draft stratified sampling guidance. *J. Pharm Innov* 2015;**10**:76–83.
28. Gershon M. Statistical process control for the pharmaceutical industry. *J Parenteral Sci Tech* 1991;**45**:41–50.
29. FDA. Guidance for Industry PAT—A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance. Available from: www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/; September 2004.
30. FDA. Innovation and Continuous Improvement in Pharmaceutical Manufacturing Pharmaceutical, cGMPs for the 21st Century.
31. Wildfong PLD, Samy AS, Corfa J, Peck GE, Morris KR. Accelerated fluid bed drying using NIR monitoring and phenomenological modeling: method assessment and formulation suitability. *J Pharm Sci* 2002;**91**(3):631–9.
32. Frake P, Greenhalgh D, Grierson SM, Hempenstall JM, Rudd DR. Process control and end-point determination of a fluid bed granulation by application of near infra-red spectroscopy. *Int J Pharm* 1997;**151**:75–80.
33. Hailey PA, Doherty P, Tapsell P, Oliver T, Aldridge PK. Automated system for the on-line monitoring of powder blending processes using near-infrared spectroscopy part I. System development and control. *J Pharm Biomed Anal* 1996;**14**:551–9.
34. Cogdill RP, Anderson CA, Delgado M, Chisholm R, Bolton R, Herkert T, et al. Process analytical technology case study: part II. Development and validation of quantitative near-infrared calibrations in support of a process analytical technology application for real-time release. *AAPS PharmSciTech* 2005;**6**(2):E273–83.
35. Spencer JA, Jefferson EH, BoClair T, Chan J: Determination of acetaminophen in tablets by near-infrared reflectance spectroscopy. In: AAPS annual meeting abstract, October, 2000.

This page intentionally left blank

CHAPTER

26

Process Development, Optimization, and Scale-Up: Providing Reliable Powder Flow and Product Uniformity

T. Baxter and J. Prescott

Jenike & Johanson, Inc., Tyngsboro, MA, United States

26.1 INTRODUCTION

Objective of the chapter

The primary focus of this chapter is to provide guidance in designing bulk solids (powder) handling equipment to provide consistent, reliable flow, and the required product uniformity by minimizing segregation. The principles discussed in this chapter can be applied to analyzing new or existing equipment designs. These principles can also be used to compare different powders, using the various test methods discussed.

The chapter will focus on the equipment used from the final blend step to the inlet of the press/encapsulation machine used to create the unit dose. This chapter is divided into these primary topics:

26.1. Introduction: A review of introductory concepts, such a flowability, blending, and segregation

26.2. Common Powder Handling Equipment: A description of the common handling equipment, and the equipment parameters that affect powder flow and product uniformity (segregation potential)

26.3. Typical Flow and Segregation Concerns: A review of common flow and segregation concerns, and the two primary flow patterns (mass flow vs funnel flow)

26.4. Measurement of Flow Properties: A summary of the flow properties that need to be measured to obtain the equipment design parameters required for consistent, reliable flow

26.5. Basic Equipment Design Techniques: A review of the basic design techniques for the blender-to-press equipment to provide reliable flow, and minimize the adverse effects of segregation

At the end of this chapter, the reader should have a working knowledge of what flow properties need to be measured, how to measure them, and how to apply them to analyze or design handling equipment for reliable flow, as well as to minimize segregation.

Motivation for the chapter

Many pharmaceutical processes include powder handling, such as blending, transfer, storage, feeding, compaction, and fluidization. A full understanding of powder flow behavior is essential when developing, optimizing, or scaling-up a process. This may include designing new equipment or developing corrective actions for existing equipment for a batch or continuous manufacturing process. There are several instances where the robustness of a process is adversely affected by flow or segregation problems that develop.

Common flow and segregation problems can have an adverse effect upon:

1. Production costs, due to reduced production rates (eg, tableting rate limitations, required operator intervention), restrictions on raw ingredient selection (eg, percentage of lubrication used, excipients selected), method of manufacturing (wet granulation vs dry granulation vs direct compression), equipment selection (type of blender, bin, press), and overall yield;

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00026-1

2. Product quality due to variation of tablet properties (weight, hardness, etc.) or segregation and content uniformity concerns; and
3. Time to market, due to delays in product/process development and failed validation or commercial batches because flow problems may not occur until the process has been scaled-up.

This chapter primarily focuses on batch processes, but provides comments on continuous manufacturing process where pertinent.

26.1.1 Introduction to flowability

A bulk solid is defined as a collection of solid particles. The term "powder" is often used to describe a fine bulk solid, especially in the pharmaceutical industry. This common term will be used predominantly throughout this chapter. The concepts discussed in this chapter apply to many types of powders with different particle sizes, shapes, and distributions. The powders may include dust, granulations, and granules. The powder could either be a single substance, such as an excipient or active pharmaceutical ingredient (API), or a multi-component blend (final blend). As the pharmaceutical industry begins to adopt more continuous manufacturing of solid oral dosages, reliable feed of the individual blend components, including the API, will be more critical. The principles outlined in this chapter can be used to design for all these different types of powders.

A simple definition of "flowability" is the ability of a powder to flow through equipment reliably. By this definition, there is often a tendency to define flowability as a single parameter of a powder, ranked on a scale from "free-flowing" to "nonflowing." Unfortunately, a single parameter is not sufficient to define a powder's complete handling characteristics. In addition, a single parameter is insufficient to fully address common handling concerns encountered by the formulator and equipment designer. In fact, several design parameters may need to be known for a successful design. The behavior of a powder will depend upon several different parameters or "flow properties." Flow properties are the specific properties of a powder that affect how flow that can be measured. Therefore, a full range of flow properties will need to be measured to fully characterize the powder. Section 26.4 discusses the measurement of these flow properties.

In addition, the "flowability" of a powder is a function of the powder's flow properties and the design parameters of the handling equipment. For example, "poor flowing" powders can be handled reliably in properly designed equipment. Conversely, "good flowing" powders may develop flow problems in improperly designed equipment. Our definition of "flowability" is, therefore, "the ability of powder to flow in the desired manner in a specific piece of equipment."

The flow properties of the powder should be quantitative and scalable design parameters. The term "flow properties" often refers to the physical characteristics of the powder that were measured. For example, one might report "the tapped density of a final blend was measured to be 0.6 grams per cubic centimeter." The term "powder flow" often refers to an observation of how the powder flowed through a given piece of equipment. For example, one might report "the powder flow through the press hopper was consistent." In discussing or reporting flowability, both the powder flow properties and the handling equipment must be included. Therefore, the measurement of the powder flow properties can be used to predict behavior in specific equipment during scale-up.

The flow properties that should be measured (Section 26.4), and how to apply the results to a reliable equipment design (Section 26.5), will be discussed in this chapter.

26.1.2 Introduction to blending

Solid blending processes are used during the manufacture of products for a wide range of industries, including pharmaceuticals. In the pharmaceutical industry, a wide range of ingredients may be blended together to create the final blend used to manufacture the solid dosage form. The range of materials that may be blended (excipients, API), presents a number of variables that must be addressed to achieve products of acceptable blend uniformity. These variables may include the particle size distribution (PSD) (including aggregates or lumps of material), particle shape (spheres, rods, cubes, plates, and irregular), presence of moisture (or other volatile compounds), particle surface properties (roughness, cohesivity), and many other variables.

The quality of the solid dosage form depends on the adequacy of the blend. Producing a uniform mixture of the drug and its excipients is paramount for being able to deliver the proper dose of the drug to the patient. Once an adequate blend is obtained, it is also critical to ensure that it does not segregate in the post-blending handling steps. Millions of dosage units may be created from a single batch, and each and every dose must be of acceptable composition to ensure the safety and efficiency of the product. Therefore, the homogeneity of pharmaceutical blends and dosage units is highly scrutinized by both manufacturers and regulatory bodies worldwide. Formulation components and process parameters involved with blending operations should be carefully selected, and validated,

to ensure uniform blends and dosage units are produced. Blend and dosage unit uniformity data is provided in regulatory submissions, and often examined during preapproval inspections. This is to ensure that blending processes produce homogeneous blends that do not segregate upon further processing into dosage units. Finally, pharmacopeias require an assessment of content uniformity to be performed on every batch of solid dosage forms manufactured.

The scale of blending operations for the preparation of pharmaceutical dosage forms ranges from the extemporaneous compounding of a few capsules by pharmacists, to large-scale production of batches containing millions of dosage units. The complexity of the blending process can vary substantially. Large-scale production batches often use equipment capable of blending hundreds of kilograms of material. Continuous blending operations require not only the mixing components within the blender itself, but also controlled feed of the incoming material, and monitoring, feedback, and control systems. Depending on the dose and characteristics of the drug substance, commercial-scale blending processes can be complex, and may require screening operations, the preparation of pre-blends, and/or the inclusion of milling operations to achieve acceptable content uniformity. Regardless of the scale of manufacture, the goal remains the same: to prepare a blend that is adequately blended, and can be further processed into dosage units that deliver the proper dose of the drug to the patient.

Blending should not be seen as an independent unit operation, but rather as an integral part of the overall manufacturing process. Blending includes producing an adequate blend, maintaining that blend through additional handling steps, and verifying that both the blend and the finished product are sufficiently homogeneous. Therefore, a complete approach should be used to assess the uniformity of blends and the subsequent dosage forms produced from them.

A review of the common types of blenders is provided in Section 26.2 and the details of scaling-up common final blending process steps (wet granulation, fluid-bed granulation, roller compaction) are discussed in other chapters.

26.1.3 Introduction to segregation

"Segregation" can be defined as having particles of similar properties (ie, size, composition, density, resiliency, static charge, etc.) preferentially distributed into different zones within given equipment or processes. Segregation most notably affects the localized concentration of the drug substance. This can result in blend and content uniformity problems. In addition, the segregation of other components of the blend can be responsible for variations in properties such as dissolution, stability, lubrication, taste, appearance, and color. Even if the blend remains chemically homogeneous, variations in particle size can affect flowability, bulk density, weight uniformity, tablet hardness, appearance, and dissolution. In addition, segregation can create concentrations of dust, which can lead to problems with agglomeration, yield, operator exposure, containment, cleanliness, and increased potential for a dust explosion. Segregation can occur in a batch or continuous manufacturing process.

Segregation can occur any time there is powder transfer, such as when discharging the final blend from the blender into a bin. Segregation can also occur when forces acting on the particles in the blend, such as air flow or vibration, are sufficient to induce particle movement. This may include handling steps upstream of a blender, including segregation of raw materials at a supplier's plant or during shipment, movement within the blender, during its discharge, or in downstream equipment. Of all of these potential instances where segregation can occur, the most common area for issues is post-blender discharge. Therefore, this chapter will focus on segregation of the final blend in the post-blending handling steps.

The current state of understanding segregation is limited to empirical descriptors of segregation mechanisms (see Section 26.4), and prior experiences with diagnosing and addressing specific segregation behaviors (Section 26.5). Unlike the flow properties tests used to assess how a powder flows though equipment, there are no current "first principle" models that adequately describe the various segregation mechanisms that may occur in a process. Therefore, one cannot currently input the particle properties of the blend components (eg, particle size and chemical composition of the excipients, and API) into a mathematical model, and obtain a prediction of segregation potential and the resulting content uniformity of the product. However, as computational models, such as discrete element modeling, continue to evolve they will able to be tuned to match specific segregation behaviors that are created in physical models to gain further insight into blending and segregation behavior. As these models evolve, they will become more powerful and have fewer assumptions and limitations. When assessing segregation concerns, it is critical to utilize as many resources as possible, such as laboratory-scale tests, to assess different segregation mechanisms (Section 26.4) and stratified blend and content uniformity data to assess potential segregation trends.

The empirical tests that can be conducted to assess the segregation potential of a final blend (Section 26.4), and how to apply the results to a reliable equipment design (Section 26.5), are discussed later in this chapter.

26.2 COMMON POWDER HANDLING EQUIPMENT

Objective of the section

The primary objective of this section is to describe the common powder handling equipment used in pharmaceutical processes. This section will also define common terms for powder handling equipment, and provide background for Section 26.3 of this chapter, which will focus upon typical flow and segregation problems.

We will review the common handling equipment and process steps that may affect the flowability and segregation of a powder. In particular, we will review the process steps for the final blend used to manufacture the dosage units commonly used in a batch process, including:

1. processing steps prior to final blending, such as milling, screening, drying, and granulation;
2. final blending;
3. discharge from the final blender;
4. intermediate bulk containers ("IBCs," "totes," bins, drums);
5. transfer from the IBC to the press (or encapsulator); and
6. feed from the hopper to the die cavity or dosator.

As continuous manufacturing in the pharmaceutical industry continues to become more common, different processing equipment may become more crucial, such as reliable feed and metering of the API, but the equipment terms outlined in this section will still often apply to understanding how powders flow. For each of these different process steps, we will review the key equipment parameters that affect the flowability and segregation of a powder. These typical handling steps serve as examples of the concerns with powder handling, but virtually any solids handling application can be analyzed in the same way.

26.2.1 Processing steps prior to final blending

It is critical to understand the physical properties of the raw ingredients (API, excipients), and how they affect the flowability and segregation of the final blend, when selecting and designing the powder handling equipment. A full understanding of the properties of the blend components is especially critical in a continuous manufacturing process. Therefore, the process steps and equipment parameters prior to the final blend step are often critical to the flowability and segregation of the final blend. There are several common pre-blending process steps that may affect the final blend, including:

1. Storage conditions of the raw ingredients, such as the temperature, relative humidity, container dimensions, and days stored at rest (inventory control) can all influence the flowability of the final blend, especially if any of the raw ingredients are hygroscopic.
2. Milling and screening steps that alter the raw ingredients', and thus the final blend's particle size, shape, and distribution. Therefore, milling and screening process parameters, such as the mill type, mill speed, screen size, mill/screen feed method (controlled vs noncontrolled feed) may all influence on the flowability and segregation potential of the final blend, especially in a direct blending process.
3. Granulation (dry roller compaction, wet granulation, fluid-bed granulation) of the API, together with select excipients, can often have a positive effect on the flowability of the final blend, especially for blends with high active loadings. The granulation parameters, especially those that influence particle size/shape/distribution, will have a significant effect on flowability and segregation potential. For roller compaction, the process parameters that dictate the PSD and shape distribution of the final blend may include the roller compactor speed, roll compactor pressure, mill type, and the screen size. The wet granulation process parameters that affect PSD and shape, as well as the moisture content, are often critical to flowability. Therefore, wet granulation parameters, such as the blade and impeller design/speed, binder addition rate, and method and identification of granulation end point, are critical to flowability, as are the granule milling conditions. Similarly, the fluid-bed granulation parameters that affect the moisture content and particle size, such as the binder addition rate/method, inlet air flow rate and temperature, drying time, end point determination (target moisture, powder temperature, exhaust air temperature), and fluidization behavior for the powder bed are all critical to flowability and segregation. The details of scaling-up different granulation processes are discussed in other chapters.
4. Pre-blending of selected raw materials, such as pre-blending a cohesive powder with a less cohesive powder to reduce the likelihood of flow problems during subsequent handling steps. Pre-blending may also be conducted to achieve a more uniform blend and reduce the segregation potential.

The measurement of flow properties and the design parameters they provide are discussed in Section 26.4,

but they can also be applied for trouble-shooting and developing corrective actions for flow problems in the pre-blending steps if needed.

26.2.2 Final blending

Final blending may be accomplished on a batch or continuous basis. In the pharmaceutical industry, batch blenders are most common, but continuous manufacturing of oral solid dosages has progressed significantly over the last decade. Therefore, this chapter focuses on different batch blending processes.

Batch blending processes consist of three sequential steps: weighing and loading the components, blending, and discharging. Unlike a continuous blender, the retention time in a batch blender is rigidly defined and controlled, and is the same for all of the particles. Batch blenders come in many different designs and sizes, and make use of a wide range of blending mechanisms.

All blenders use one or more of the following mechanisms to induce powder blending: convective blending, shear blending, and diffusive blending.[1,2] Another classification system for blenders is based on their design. This system categorizes blenders into two categories:

- those that achieve blending within a moving vessel; and
- those with fixed vessels that rely on internal paddles or blades to move the materials.

There are many common blenders, or other techniques for making a uniform material, used in the pharmaceutical industry on a batch basis, including:

1. wet granulators;
2. fluid-bed granulators;
3. roller compactors; and
4. direct blending via tumble blenders (reviewed briefly in this chapter).

Many of these blending processes, including high-shear wet granulation, fluid-bed granulation, roller compaction, and direct blending in mixer can also be conducted in a continuous manufacturing process. There are other types of blenders, such as pneumatic blenders, extruders, ribbon blenders, planetary blenders, and orbiting screw blenders, which are not discussed in this chapter.

As a result of the multiple classification systems, a number of terms have evolved throughout the industry to describe families of blenders. Regardless of the terminology used to classify the blender, it is important for the pharmaceutical scientist to understand the capabilities and limitations of the equipment when selecting an appropriate blender for a particular product. This is especially important during process scale-up, when equipment with a different design and operating principle may need to be used. The optimization and scale-up of the wet granulation, fluid-bed granulation, and roller compaction processes are discussed in other chapters. The optimization and scale-up of tumble blenders is discussed in the literature.[3]

Note that some types of tumble blenders, such as bin blenders, serve a dual purpose. In addition to providing the container in which blending is accomplished, bin blenders can also be used to transfer the powder blend to the next unit operation in a batch process. This is of particular value when manufacturing blends that tend to segregate when discharging the blend onto the compression or filling equipment. This also makes the use of bin blenders desirable during the manufacture of potent drug products that must be processed in high containment facilities. In addition, by decoupling the blending bin from the drive mechanism, the bin filling, discharge, and cleaning take place at a separate time and location, which increases the efficiency and utilization of equipment.

26.2.2.1 Discharge from a blender or processing vessel

Powder that has been blended must be discharged from the blender for further processing, including the creation of the unit dose at the press or encapsulator. In many dry blending processes, such as tumble blending, the discharge is driven by gravity alone. As an example, the final blend step may be conducted in a V-blender or a double-cone blender. In these cases, the blender geometry often consists of a converging cross-section to the outlet, through which the powder must be discharged reliably. In these cases, the blender is essentially acting as a "bin," so the equipment parameters of interest are those that are crucial to a bin design. These crucial bin design parameters are discussed in the next section on IBCs.

For fluid-bed granulation processes, it is not uncommon for a conical "hopper" to be attached to the "bowl" of a fluid-bed granulator, inverted, and discharged to a downstream process step via gravity.

For wet granulators, the final blend may be discharged using mechanical agitation by continuing to operate the plow blade (typically at a lower speed) to discharge the final blend through a central or side outlet. Although the plow blade typically ensures that the blend is discharged from the granulator "bowl" reliably, the design of the transition chute from the blender to the downstream process is also critical. This is especially critical if the equipment below the

granulator has a converging cross-section that is full of material, as discussed further in Section 26.5.

When a blender/vessel is discharged manually (hand-scooping), flowability may not be a primary concern, but segregation concerns and other factors (eg, production rate concerns, operator exposure and safety) may limit the extent to which a blender/vessel can be manually unloaded.

The transfer of the material from the final blender to the downstream equipment, whether it is into an IBC or directly to the press, is critical to the segregation potential. Therefore, parameters such as the transfer method (eg, manual, gravity, pneumatic), transfer rate, transfer chute height and geometry, venting, and other items will be critical in assessing and minimizing the segregation potential during the blender-to-press transfer steps. Design techniques to minimize segregation during these transfer steps are further discussed in Section 26.5.

26.2.3 Intermediate bulk containers

The flowability and segregation potential of the final blend is especially critical during storage and discharge from an IBC. The IBC may be a bin (tote) or a drum that is used to store and transfer the final blend from the blender to the press. When a drum is used, an attachment such as a conical hopper may be attached to the cone to connect the drum to downstream equipment with a smaller inlet (eg, press hopper). This section focuses on the use of IBCs to store the blend in a batch process, but the concepts can also be applied and are critical to the storage of blend components in a continuous process.

An IBC consists of two primary sections (see Fig. 26.1):

1. A cylinder or straight-sided section with a constant cross-sectional area that is often rectangular (with or without radiused corners) or circular;
2. A hopper section with a changing cross-sectional area that is often a converging conical or pyramidal hopper.

IBCs may be used to store the blend for extended periods of time. During this time, the flowability may deteriorate as the blend is subjected to consolidation pressures due to its own weight during storage at rest. In addition, IBCs may be used to move the blend from one process step to another, during which time the blend may be subjected to vibration

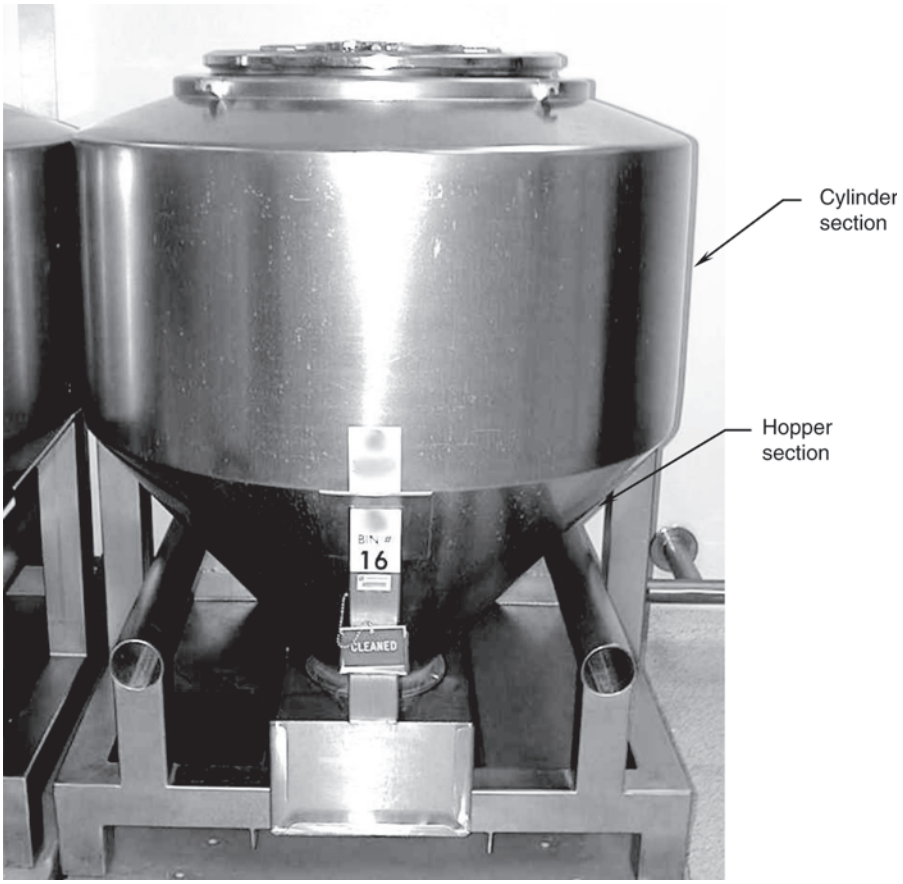


FIGURE 26.1 Intermediate bulk container ("IBC," "bin," "tote").

that may adversely affect flowability. Therefore, it is important to determine what consolidation pressures will act on the powder as it is stored and transferred in an IBC.

The key IBC equipment parameters with respect to flowability and segregation potential include:

1. The cylinder cross-sectional area and height which, along with other parameters such as the fill height, will affect the consolidation pressure acting on the blend.
2. The hopper geometry (eg, planar vs circular) and angles, which will affect the flow pattern that develops during discharge. The flow pattern, discussed in Section 26.3, will also affect the segregation potential.
3. The interior surface finish of the hopper section, which will affect the flow pattern that develops during discharge.
4. The IBC outlet size and shape (slotted vs circular), which will affect whether the blend will discharge reliably without arching or ratholing.
5. General flow impediments, such as upward facing ledges or partially opened valves, that may act as flow obstructions.

Section 26.4 discusses the measurement of the flow properties and segregation potential that are used to obtain the key bin design parameters. The application of these design parameters to provide reliable flow and minimize segregation from the IBC to the press is discussed in Section 26.5.

26.2.3.1 Transfer from intermediate bulk containers to the press/encapsulator

The flowability and segregation potential of the final blend is also critical during transfer from the IBC to the press/encapsulation machine/etc. This transfer step may be a manual transfer (hand-scooping), in which case flowability may not be a primary concern. The transfer step may also be conducted via pneumatic conveying, in which case the flowability of the blend may not be a primary concern, but equipment and material parameters affecting conveying (conveying gas pressure and flow rate, conveying line diameter and layout, etc.) need to be considered. Pneumatically conveying the final blend may also raise segregation concerns, as further discussed in Section 26.3.

The transfer step may also be conducted via gravity transfer via a single or bifurcated chute (see Fig. 26.2), depending on the press configuration. These chutes are often operated in a flood-loaded manner, and may consist of converging sections where the cross-sectional area of the chute is reduced. If this is the case, the chute will need to be designed for reliable flow in a similar manner to the IBCs.

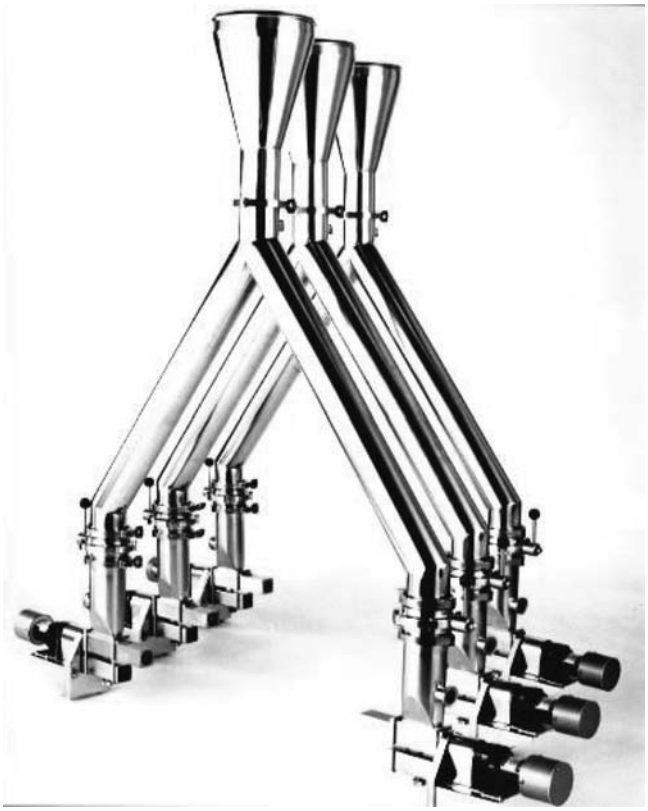


FIGURE 26.2 Bifurcated press feed chute.

The key transfer chute parameters with respect to flowability and segregation potential include:

1. The chute cross-sectional area and height, which will affect the consolidation pressure acting on the blend, and how it may segregate.
2. Valving and venting of the transfer chute may affect how readily a blend segregates.
3. For converging and nonconverging sections of the chute, the chute geometry, angles, and interior surface finish will affect the flow pattern that develops during discharge through the chute.
4. For converging sections of the chute, the outlet shape and size will affect whether the blend will discharge reliably without arching or ratholing.
5. General flow impediments, such as upward facing ledges (mismatched flanges), sight glasses, level probes, or partially opened valves may act as flow obstructions.

The measurement of the flow properties and segregation potential used to obtain the key chute design parameters are discussed in Section 26.4. The application of these design parameters for equipment from the IBC to the press is discussed in Section 26.5.

26.2.3.2 *Feed from the press hopper to the die cavity*

The press feed hopper must also be designed to provide reliable flow, and minimize segregation. Most presses utilize a small press hopper that is, in essence, a miniature IBC designed to provide a small amount of surge capacity. The press hopper often consists of a cylinder section and a hopper section similar to a larger IBC. The hopper section may be asymmetric, as opposed to the symmetric hopper designs commonly used for larger IBCs. The press hopper is typically flood-loaded from the IBC/chute above via gravity feed. However, in some instances, the material level in the press hopper may be controlled via a feeder at the IBC outlet (eg, rotary valve or screw feeder), and/or use agitation to promote flow. Some modern presses do not have press hoppers with a converging hopper, but instead consist of vertical, nonconverging chutes from the press inlet to the feed frame inlet.

The key equipment parameters with respect to flowability, which are outlined in the preceding section for IBCs/bins, are also applicable to the press hopper. Since the press hopper outlets are often much smaller than an IBC outlet, flow problems, such as arching or ratholing (discussed further in Section 26.3), may be more pronounced at this location. As a result, press hoppers may include mechanical agitators to assist gravity discharge, such as a rotating agitator mounted to a vertical shaft (Fig. 26.3).

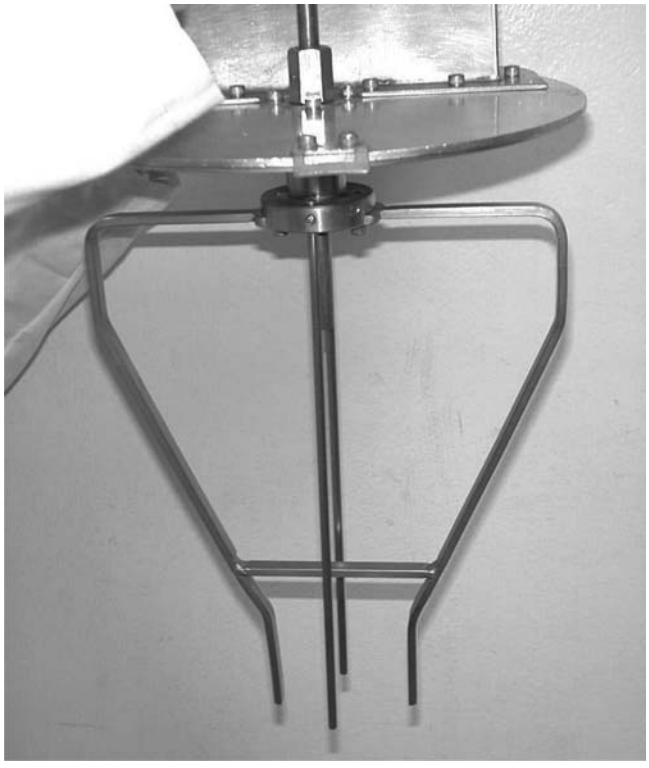


FIGURE 26.3 Agitator for use in small-scale hopper.

The same design parameters used for a reliable IBC design (see Section 26.4) can also be used to design a press hopper (see Section 26.5).

26.3 TYPICAL FLOW AND SEGREGATION CONCERNS

Objective of the section

This section will focus on typical flow and segregation concerns that occur during transfer operations from the final blender to the press. This section will also provide a summary of common flow problems, segregation mechanisms, and the flow patterns that can occur during gravity discharge.

26.3.1 Common flow problems

A number of problems can develop as powder flows through equipment such as bins, chutes, and press hoppers. If the powder is cohesive, an arch or rathole may form, resulting in "no flow" or erratic flow. In addition, flooding or uncontrolled discharge may occur if a rathole spontaneously collapses. A deaerated bed of fine powder may experience flow rate limitations or no flow conditions due to the two-phase flow effects between the powder and the interstitial air. These flow problems are primarily reviewed with respect to batch processes, but also commonly apply to continuous processes. Each of these flow problems is discussed in more detail next.

- No flow: no flow from a bin/hopper is a common and significant solids-handling problem, especially when handling fine/milled APIs and high active loading blends. In production, it can result in problems, such as starving downstream equipment, production delays, and the requirement for frequent operator intervention to reinitiate flow. No flow can be due to either arching (sometime referred to as "bridging" or "plugging") or ratholing (also referred to as "piping") as described next.
  - Arching: occurs when an obstruction in the shape of an arch or bridge forms above the bin outlet, and prevents any further material discharge. It can be an interlocking arch, where the particles mechanically lock to form the obstruction, although this is less common with fine pharmaceutical powders. An interlocking arch occurs when the particles are large compared with the outlet size of the hopper. The arch could also be a cohesive arch where the particles pack together to form an obstruction (Fig. 26.4). Both of these problems are strongly influenced by the outlet size of the hopper the material is being fed through. Powder flow

FIGURE 26.4 Cohesive arch.

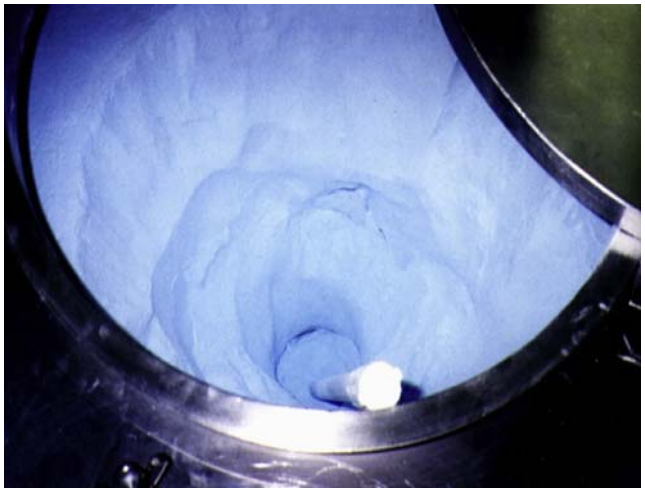


FIGURE 26.5 Rathole.

properties, discussed in Section 26.4, can be used to determine if these problems will occur and used to address them during scale-up. In particular, the cohesive strength of a powder will dictate what size outlet it can arch over (the greater the cohesive strength, the higher the likelihood of arching).

- Ratholing: can occur in a bin when the powder empties through a central flow channel, but the material at the bin walls remains stagnant and leaves an empty hole (rathole) through the material, starting at the bin outlet (Fig. 26.5). Ratholing is influenced by the bin/hopper geometry and outlet size the material is fed through. Similar to the problem of arching, this problem will arise if the material has sufficient cohesive strength. In this case, the material discharge will stop once the flow channel empties.
- Erratic flow: erratic flow is the result of obstructions alternating between an arch and a rathole. A rathole may collapse due to an external force, such as vibrations created by surrounding equipment or a flow-aid device, such as an external vibrator. While some material may discharge, falling material may impact over the outlet and form an arch. This arch may then break, due to a reoccurrence of the external force, and material flow may not resume until the flow channel is emptied and a rathole is formed again. This not only results in erratic feed to the downstream equipment (press), but can also result in a nonuniform feed density.
- Fine powder flow concerns (two-phase flow effects): additional flow concerns can arise when handling fine powders, generally in the range below 100 μm in average particle size. These concerns are due to the interaction of the material with entrained air or gas, which becomes significant in describing the behavior of the material. This interaction can result in two-phase (powder/interstitial gas) flow effects. There are three modes that can occur when handling fine powders that are susceptible to two-phase flow effects: steady flow, flooding (or flushing), and a flow rate limitation.[4] These three flow modes are discussed in more detail next.
  - Steady flow: will occur with fine powders if the target flow rate (feed rate through the system) is below the "critical flow rate" that occurs when the solids stress is balanced by the air pressure at the outlet. The target flow rate is often controlled by a feeder, such as at the inlet to a compression machine (press feed frame). The critical flow rate and the flow properties tests used to determine it are described in more detail in Section 26.4. At target flow rates exceeding the critical flow rate, unsteady flow can occur by two different modes:
    - Flooding: (or "flushing") is an unsteady two-phase flow mode that can occur as falling particles entrain air and become fluidized. Since powder handling equipment often cannot contain a fluid-like powder, powder can flood through the equipment (feeders, seals) uncontrollably. Flooding can also occur when handling fine powders in small hoppers with high fill and discharge rates. In such situations, the powder does not have sufficient residence time to deaerate, resulting in flooding through the feeder. One adverse effect of flooding or flushing may be high variation in the tablet weight and strength.

– Flow rate limitation: is another unsteady two-phase flow mode that can occur with fine powders. Fine powders have very low permeabilities, and are affected by any movement of the interstitial air (air between the particles). This air movement will occur due to the natural consolidation and dilation of the powder bed that takes place as it flows through the cylindrical and hopper geometries. As the material is consolidated in the cylinder, the air is squeezed out. As the powder flows through the hopper and outlet, it dilates and additional air must be drawn in. The air pressure gradients caused as a result of this air movement can retard discharge from a hopper, significantly limiting the maximum achievable rates. This may be observed when the speed of a high-speed press is increased, and the tablet weight variation increases.

During unsteady two-phase flow modes, the material's bulk density can undergo dramatic variations. This can negatively impact downstream packaging or processing operations. Problems can result, such as excessive tablet weight variation, a required reduction in filling speeds, and even segregation. Equipment and process parameters will govern whether such problems occur, and are further discussed in Section 26.5. These parameters include hopper geometry and outlet size, applied vacuum and other sources of air pressure differences (such as dust collection systems), material level, time since filling, and of course the target feed rate.

Material properties, such as permeability and compressibility (discussed in Section 26.4) will also play important roles, as will variations in the material's state of aeration that can occur based on its residence time or degree of compaction from external forces and handling.

One of the most important factors in determining whether a powder will discharge reliably from a hopper is establishing what flow pattern will develop, which is discussed in the next section.

26.3.2 Flow patterns

Two flow patterns can develop in a bin or hopper: funnel flow and mass flow. In funnel flow (Fig. 26.6), an active flow channel forms above the outlet, which is surrounded by stagnant material. This results in a first-in, last-out flow sequence. It generally occurs in equipment with relatively shallow hoppers. Common examples of funnel flow bins include hopper geometries, such as asymmetric cones and rectangular-to-round transitions (Fig. 26.7).

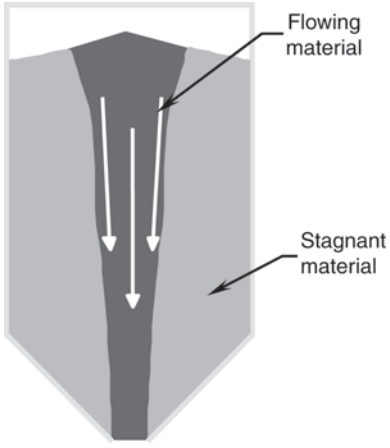


FIGURE 26.6 Funnel flow discharge pattern.

FIGURE 26.7 Examples of funnel flow bins.

FIGURE 26.8 Mass flow discharge pattern.

As the level of powder decreases in funnel flow, stagnant powder may fall into the flow channel if the material is sufficiently free-flowing. If the powder is cohesive, a stable rathole may develop. Funnel flow occurs if the powder is unable to flow along the hopper walls, due to the combination of friction against the walls and hopper angle. Funnel flow's first-in-last-out flow sequence can also have an adverse effect upon segregation, and result in a nonuniform feed density to the downstream equipment (tablet weight variation).

In mass flow (Fig. 26.8), all of the powder is in motion whenever any is withdrawn. Powder flow occurs throughout the bin, including at the hopper walls. Mass flow provides a first-in-first-out flow sequence, eliminates stagnant powder, provides a steady discharge with a consistent bulk density, and provides a flow that is uniform and well-controlled. Ratholing will not occur in mass flow because all of the material is in motion. A mass flow bin design may also be beneficial with respect to reducing segregation (further discussed in Section 26.5).

The requirements for achieving mass flow are:

1. sizing the outlet large enough to prevent arch formation; and
2. ensuring the hopper walls are steep and smooth enough to allow the powder to flow along them.

These mass flow design parameters are discussed further in Section 26.4, and their implementation is discussed in Section 26.5.

26.3.3 Common segregation mechanisms

The term "segregation mechanism" refers to the mode or motive force by which the components of the blend separate. There are many different segregation mechanisms that can adversely affect uniformity,[5] but three primary segregation mechanisms are of interest in typical pharmaceutical blend handling operations. These three primary segregation mechanisms are:

1. sifting segregation (sometimes referred to as "percolation segregation");
2. fluidization segregation (sometimes referred to as "air entrainment"); and
3. dusting segregation (sometimes referred to as "particle entrainment in an air stream").

The three segregation mechanisms are described in more detail next. These terms are not universally defined, so one must use caution when using them. Segregation may occur as a result of just one of these mechanisms or a combination of several mechanisms.

26.3.3.1 Material properties that affect segregation

Whether segregation occurs, to what degree, and which mechanism or mechanisms are involved depends on a combination of the properties of the blend, and the process conditions encountered. Some of the primary material properties that influence segregation tendencies include:

1. The mean particle size and PSD of the API, excipients, and final blend: segregation can occur with blends of any mean size, but different mechanisms become more pronounced at different particle sizes, as further discussed next.
2. Particle density: the particle density will affect how the blend components fluidize.
3. Particle shape: rounded particles may have greater mobility than irregularly shaped particles, which can allow more segregation.
4. Particle resilience: this property influences collisions between particles and surfaces, which can lead to differences in where components accumulate during the filling of a bin or press.
5. Cohesive strength of the blend: as a general rule, more cohesive blends are less likely to segregate. However, if enough energy is added to dilate the blend and/or separate particles from one another, even a very cohesive material can segregate.
6. Electrostatic effects: the ability of components to develop and hold an electrostatic charge, and their affinity for other ingredients or processing surfaces, can also contribute to segregation tendencies.

Of all of these, segregation based on particle size is by far the most common.[6] In fact, particle size is the most important factor in the three primary segregation mechanisms considered here, as further described in the following sections.

26.3.3.2 Sifting segregation

Sifting segregation is common for many industrial processes. Under appropriate conditions, fine particles tend to sift or percolate through coarse particles. For segregation to occur by this mechanism, four conditions must exist:

1. There must be a range of particle sizes. A minimum difference in mean particle diameters between components of 1.3:1 is often more than sufficient.[6]
2. The mean particle size of the mixture must be sufficiently large, typically greater than approximately 100 μm.[7]
3. The mixture must be relatively free-flowing to allow particle mobility.
4. There must be relative motion between particles (interparticle motion). This last requirement is very important because without it even highly-segregating blends of ingredients that meet the first three tests will not segregate. Relative motion can be induced in a variety of ways, such as when a pile is formed when filling a bin, vibration from surrounding equipment (such as a tablet press), or as particles tumble and slide down a chute.

If any one of these conditions does not exist, the mix will not segregate by this mechanism.

The result of sifting segregation in a bin is usually a side-to-side variation in the PSD. The smaller particles will generally concentrate under the fill point, with the coarse particles concentrating at the perimeter of the pile (Fig. 26.9).

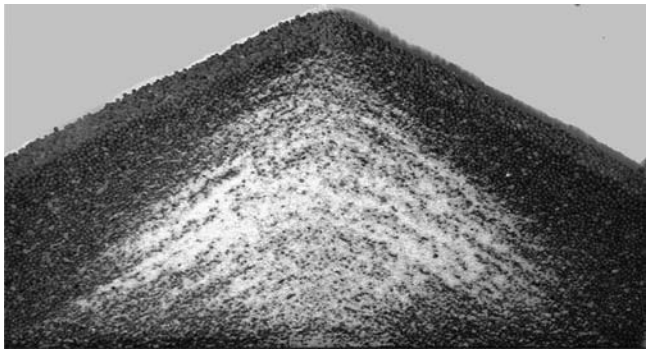


FIGURE 26.9 Example of sifting segregation in a 2-D pile. Note the black particles are approximately 1200 μm, while the white particles are approximately 350 μm. *Source: Photo provided courtesy of Jenike & Johanson, Inc.*

26.3.3.3 Fluidization segregation

Variations in particle size or density often result in vertically segregated material when handling powders that can be fluidized. Finer or lighter particles often will be concentrated above larger or denser particles. This can occur during filling of a bin or other vessel or within a blending vessel once the blending action has ceased.

Fluidization segregation often results in horizontal gradation of fine and coarse material. A fine powder can remain fluidized for an extended period of time after filling or blending. In this fluidized state, larger and/or denser particles tend to settle to the bottom. Fine particles may be carried to the surface with escaping air as the bed of material deaerates. For example, when a bin is being filled quickly, the coarse particles move downward through the aerated bed, while the fine particles remain fluidized near the surface. This can also occur after blending, if the material is fluidized during blending.

Fluidization is common in materials that contain a significant percentage of particles smaller than 100 μm.[8] Fluidization segregation is most likely to occur when fine materials are pneumatically conveyed, when they are filled or discharged at high rates, or if gas counterflow occurs. As with most segregation mechanisms, the more cohesive the material, the less likely it will be to segregate by this mechanism.

Fluidization via gas counterflow can occur as a result of insufficient venting during material transfer. For example, consider a tumble blender discharging material to a drum, with an airtight seal between the two (Fig. 26.10). As the blend transfers from the blender to the drum, air in the drum is displaced, and a slight vacuum is created in the blender. If both are properly vented, air moves out of the drum and, separately, into the blender. If the bin and drum are not vented, the air must move from the drum to the blender through the blender discharge. In doing so, the fines may be stripped off of the blend, and carried to the surface of the material still within the blender.

26.3.3.4 Dusting segregation

Like fluidization segregation, dusting is most likely to be a problem when handling fine, free-flowing powders with particles smaller than about 50 μm,[8] as well as a range of other particle sizes. If dust is created upon filling a bin, air currents created by the falling stream will carry particles away from the fill point (Fig. 26.11). The rate at which the dust settles is governed by the particle's settling velocity. The particle diameter is much more significant than particle density in determining settling velocity.

As an example of this mechanism, consider a mix of fine and large particles that is allowed to fall into the

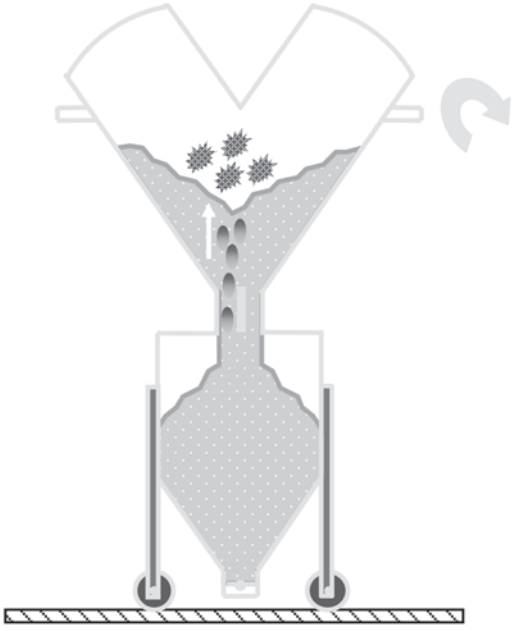


FIGURE 26.10 Example of fluidization segregation from blender to bin.

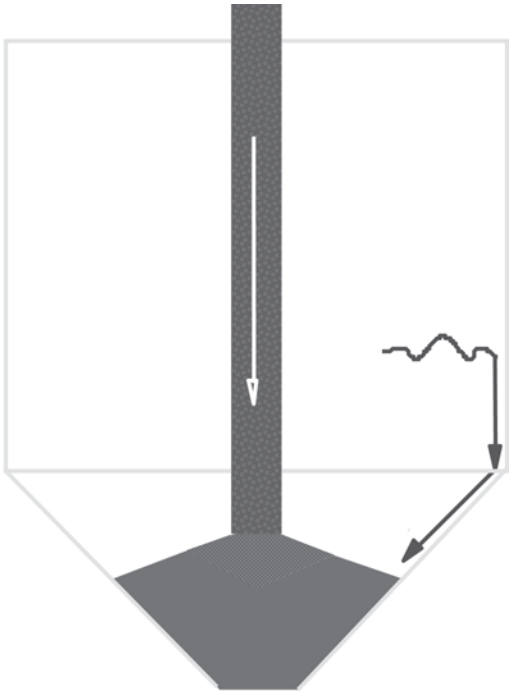


FIGURE 26.11 Example of dusting segregation during filling of bin.

center of a bin. When the stream hits the pile of material in the bin, the column of air moving with it is deflected. The air then sweeps off the pile toward the perimeter of the bin, where the air becomes highly disturbed. After, the air generally moves back up the bin walls in a swirling pattern. At this point, the gas velocity is much lower, allowing many particles to fall out of suspension. The finest particles (with low settling velocities) will be carried to the perimeter of the bin because settling velocity is a strong function of particle diameter. The larger particles will concentrate closer to the fill point where the air currents are strong enough to prevent the fine particles from settling. Dusting segregation can also result in less predictable segregation patterns, depending on how the bin is loaded, venting in the bin, and dust collection use and location.

The empirical tests used to assess these segregation mechanisms are reviewed in Section 26.4, and the application of their results to design equipment to minimize segregation is presented in Section 26.5.

26.4 MEASUREMENT OF FLOW PROPERTIES

Objective of the section

The primary objective of this section is to provide a description of the different quantitative flow properties that should be measured, as well as the empirical segregation test methods used to assess segregation potential by different mechanisms. This section also reviews the calculation of equipment parameters required for consistent, reliable flow in powder handling equipment for gravity discharge.

The following primary topics are reviewed in this section:

1. Cohesive strength tests: a review of different shear test methods, the Jenike Direct Shear Test method, and the calculation of the design parameters to prevent arching and ratholing.
2. Wall friction tests: a review of the Jenike Direct Shear Test method for measuring wall friction, and the calculation of the design parameters to provide mass flow (mass flow hopper angles).
3. Bulk density test: a review of different bulk density test methods, the compressibility test method, and the application of the compressibility test results.
4. Permeability: a review of the permeability test method, and application of the results (critical flow rate).
5. Segregation tests: a review of two empirical test methods to assess segregation potential.

Note that additional powder characterization methods are discussed in chapter "Fundamental of Diffusion and Dissolution."

26.4.1 Cohesive strength tests: preventing arching and ratholing

Dr Andrew Jenike developed his mathematical model of the flow of powders by modeling the powder as a rigid-plastic (not a visco-elastic) continuum of solid particles.[9] This approach included the postulation of a "flow–no flow" criterion that states the powder would flow (eg, from a bin) when the stresses applied to the powder exceeded the strength of the powder. The strength of a material will vary, depending on how consolidated it is. For example, the strength of wet sand increases as the consolidation pressure is increased (eg, the more one packs it). Therefore, it is critical to be able to measure the cohesive strength of a powder when scaling-up a process and designing equipment.

26.4.1.1 Test methods

One of the primary flow problems that can develop in powder handling equipment is a no-flow obstruction due to the formation of a cohesive arch or rathole, as discussed in Section 26.3. The required outlet size to prevent a stable cohesive arch or rathole from forming is determined by applying the flow-no flow criterion, and using the results of a cohesive strength test. In order to apply the flow-no flow criterion we need to determine:

1. The cohesive strength of the material as a function of the major consolidation pressure acting on the material: the consolidation pressure acting on the powder changes throughout the bin height due to the weight of material above it. Therefore, the cohesive strength must be measured over a range of consolidation pressures. The cohesive strength can be measured as a function of major consolidating pressure using the test methods described in this section.
2. The stresses acting on the material to induce flow: gravity pulls downward on a potential arch that may form. The stresses acting on the powder can be determined using mathematical models.[9]

To further illustrate the concepts of strength and consolidation pressure, consider an "idealized" strength test, as shown in Fig. 26.12. In this idealized test, the cohesive strength of the powder is measured in the following distinct steps:

1. Consolidation of the powder: the powder is consolidated using a prescribed consolidation pressure (P). In the idealized test shown, the sample is contained in a cylinder and the consolidation pressure is applied from the top (Step 1 in Fig. 26.12).
2. Fracture of the powder: once the consolidation pressure is applied, the cylinder containing the powder would be removed without disturbing the powder sample (Step 2a in Fig. 26.12). After, the strength of the powder can be measured by applying pressure to the column of powder until it fails or fractures (Step 2b in Fig. 26.12). The applied pressure at which the powder failed is referred to as the yield strength (F) (cohesive strength). This idealized test could be repeated several times to develop a flow function (FF), which is a curve illustrating the relationship between the yield strength (F) and the major consolidation pressure (P). An example of a flow function is shown in Fig. 26.12.

This idealized strength test is not possible for the broad range of powders that might be tested.

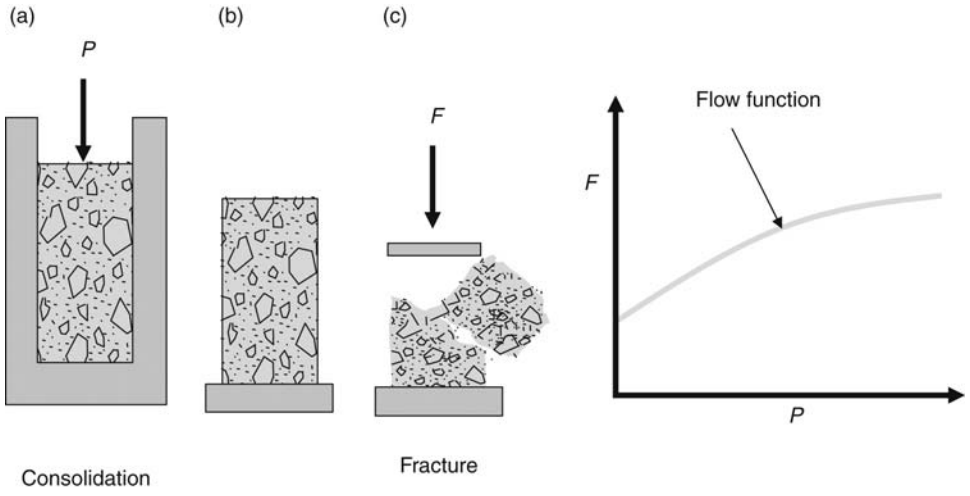


FIGURE 26.12 Schematic of "idealized" strength test: (a) Step #1; (b) Step #2a; (c) Step #2b.

Therefore, several different cohesive strength test methods have been developed. The respective strengths and weaknesses of these different cohesive strength tests are reviewed in the literature.[10,11] Although many different test methods can be used to measure cohesive strength, this chapter focuses specifically upon the Jenike Direct Shear Test method. The Jenike Direct Shear Test method is the most universally accepted method, and is described in ASTM standard D 6128.[12] It is important that these tests be conducted at representative handling conditions, such as temperature, relative humidity, and storage at rest, because all these factors can affect the cohesive strength. An arrangement of a cell used for the Jenike Direct Shear Test is shown in Fig. 26.13. The details of the Jenike Direct Test Method are provided in Jenike,[9] including the generation of:

1. Mohr's circle to plot the shear stress (τ) versus the consolidation pressure (σ);
2. Effective yield locus; and
3. Flow function.

The data generated experimentally from the Jenike Direct Shear Test can be used to determine the following derived parameters:

1. The flow function that describes the cohesive strength (unconfined yield strength, f_c) of the powder as a function of the major consolidating pressure (σ_1). The flow function is one of the primary parameters used to calculate the minimum outlet diameter/width for bins, press hoppers, blender outlets, and others, to prevent arching and ratholing. The calculation of the minimum outlet diameter/width is discussed in more detail next.
2. The effective angle of internal friction (δ) that is used to calculate the minimum outlet to prevent arching and the required hopper angles for mass flow (described next).
3. The angle of internal friction for continuous flow or after storage at rest (ϕ and ϕ_t). The static angle of internal friction (ϕ_t) is used to calculate the minimum outlet to prevent ratholing (described in the next section).

There are other testing methods that utilize the same principles of consolidation and shearing to determine the cohesive strength of a bulk powder. Annular (ring) shear testers produce rotational displacement between cell halves containing material, rather than a lateral displacement. The loading and shearing operations are more readily adapted to automation, since unlimited travel can be achieved with this type of test cell. The successful use of an annular ring shear tester to measure cohesive strength has been discussed in the industry.[13–16]

26.4.1.2 Calculation of minimum required outlet dimensions to prevent arching (mass flow bin)

The flow behavior of powders through bins and hoppers can be predicted by a complete mathematical relationship. This is very beneficial when scaling-up a process, and designing powder handling equipment. If gravity discharge is used, the outlet size required to prevent a cohesive arch or rathole from forming over a bin outlet can be calculated. The term B_c refers to the minimum outlet diameter for a circular outlet

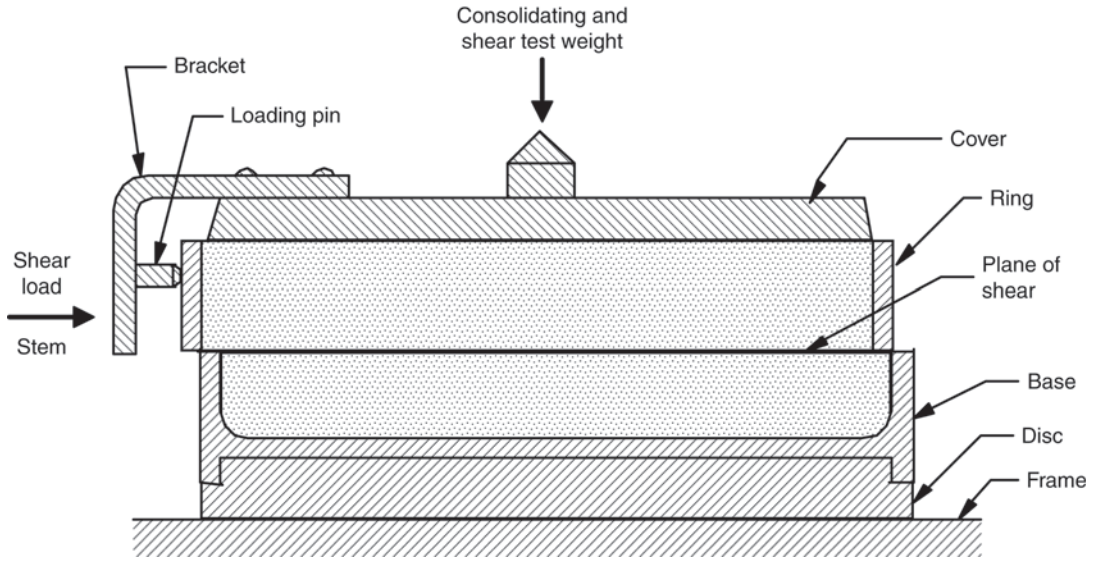


FIGURE 26.13 Jenike Direct Shear Test, cohesive strength test set-up.

to prevent a cohesive arch from forming in a mass flow bin. The term B_p refers to the minimum outlet width for a slotted outlet (B_p), in which the length: width ratio exceeds 3:1, to prevent arching in a mass flow bin.

The majority of bins used in pharmaceutical processes utilize hoppers with circular outlets. Therefore, we will focus our discussion on the calculation of the B_c parameter. It is worth noting that the outlet diameter required to prevent arching over a circular outlet (B_c) will typically be about two times greater than the required outlet width of a slotted outlet (B_p). The calculation of B_p is provided in Jenike.[9]

For mass flow, the required minimum outlet diameter to prevent arching is calculated in Eq. (26.1):

$$B_c = H(\theta')f_{crit}/\gamma \tag{26.1}$$

where:

$H(\theta')$ is a dimensionless function derived from first principles (mathematical model) and is given by Fig. 26.14. The complete derivation of $H(\theta')$ is beyond the scope of this chapter, but is provided in Jenike.[9]

f_{crit} (units of force/area) is the unconfined yield strength at the intersection of the hopper flow factor (ff) and the experimentally derived flow function (FF), as shown in Fig. 26.15. The flow factor (ff) is a mathematically determined value that represents the minimum available stress available to break an arch. The calculation of the flow factor (ff) is also beyond the scope of this chapter, but is provided in Jenike[9] and is a function of the flow properties and the hopper angle (θ).

γ (units of weight/volume) is the bulk density of the powder at the outlet.

This calculation yields a dimensional value of B_c in units of length, which is scale-independent. Therefore, for a mass flow bin, the opening size required to prevent arching is not a function of the diameter of the bin, height of the bin, or the height-to-diameter ratio.

The determination of B_c is especially valuable in making decisions during process scale-up. As a formulation is developed, a cohesive strength test can be conducted early in the development process to determine the cohesive strength (flow function). This material-dependent flow function, in conjunction with Eq. (26.1), will yield a minimum opening (outlet) size to prevent arching in a mass flow bin. For example, this opening size may be calculated to be 200 mm. This 200 mm diameter will be required whether the bin holds 10 or 1000 kilos of powder, and is scale-independent. In this example, feeding this material through a press hopper or similarly small openings would pose problems with an arch developing over the outlet because a 200 mm diameter opening is required. This information could then be used early in the development process to consider reformulating the product to reduce the cohesive strength, and improve flowability.

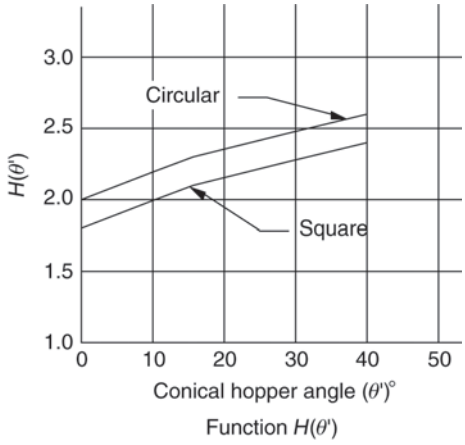


FIGURE 26.14 Plot of derived function $H(\theta')$ used to calculate arching dimensions for mass flow bins.

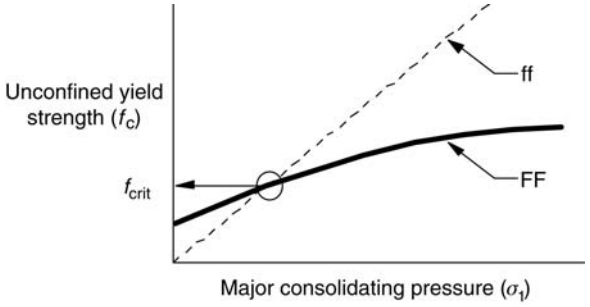


FIGURE 26.15 Example of flow function (FF) and flow factor (ff) intersection, showing f_{crit} at their intersection.

26.4.1.3 Calculation of minimum required outlet dimensions to prevent ratholing (funnel flow bin)

If the bin discharges in funnel flow, the bin outlet diameter should be sized to be larger than the critical rathole diameter (D_f) to prevent a stable rathole from forming over the outlet. For a funnel flow bin with a circular outlet, sizing the outlet diameter to exceed the D_f will also ensure that a stable arch will not form (because a rathole is inherently stronger than an arch). The D_f value is calculated in Eq. (26.2), and additional details of the calculation are provided in Jenike.[9]

$$D_f = G(\phi_t)f_c(\sigma_1)/\gamma \tag{26.2}$$

where:

$G(\phi_t)$ is a mathematically derived function from first principles and is given by Fig. 26.16.

$f_c(\sigma_1)$ is the unconfined yield strength of the material. This value is determined by the flow function (FF) at the actual consolidating pressure σ_1 (discussed next).

The consolidation pressure σ_1 is a function of the "head" or height of powder above the outlet of the bin, and takes into account the load taken up by friction along the walls, as derived by Janssen[17] in Eq. (26.3):

$$\sigma_1 = (\gamma R/\mu k)\ (1 - e^{(-\mu k h/R)}) \qquad (26.3)$$

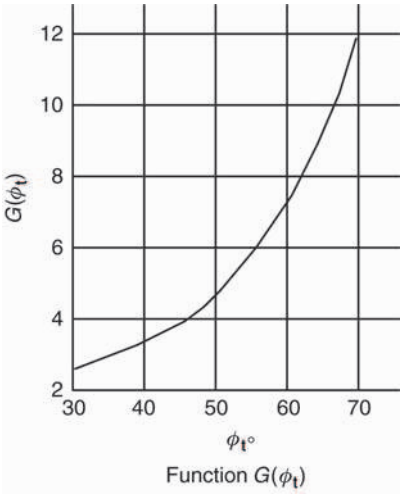


FIGURE 26.16 Plot of derived function $G(\phi_t)$ used to calculate critical rathole diameter for funnel flow bins.

where:

γ is the average bulk density of the powder in the bin
R is the hydraulic radius (area/perimeter)
μ is the coefficient of friction, (μ = tangent ϕ'). The value of ϕ' is determined from the wall friction test (discussed in the next section)
k is the ratio of horizontal to vertical pressures. A k value of 0.4 is typically used for a straight sided section.
h is the depth of the bed of powder within the bin.

This relationship in Eq. (26.2) cannot be reduced further (eg, to a dimensionless ratio) because the function $f_c(\sigma_1)$ is highly material-dependent.

The application of these parameters (B_c, D_f) to design new equipment or develop corrective equipment modifications is further discussed in Section 26.5.

26.4.1.4 *Wall friction: determining hopper angles for mass flow*

Test Method: The wall friction test is crucial in determining whether a given bin will discharge in mass flow or funnel flow (mass flow and funnel flow are discussed in the previous section). Wall friction is caused by the powder particles sliding along the wall surface. In Jenike's continuum model,[9] wall friction is expressed as the wall friction angle (ϕ'). The lower the wall friction angle, the shallower the hopper or chute walls need to be for powder to flow along them. This coefficient of friction can be measured by shearing a sample of powder in a test cell across a stationary wall surface using a Jenike Direct Shear tester.[9,12] One arrangement of a cell used for the wall friction test is shown in Fig. 26.17. In

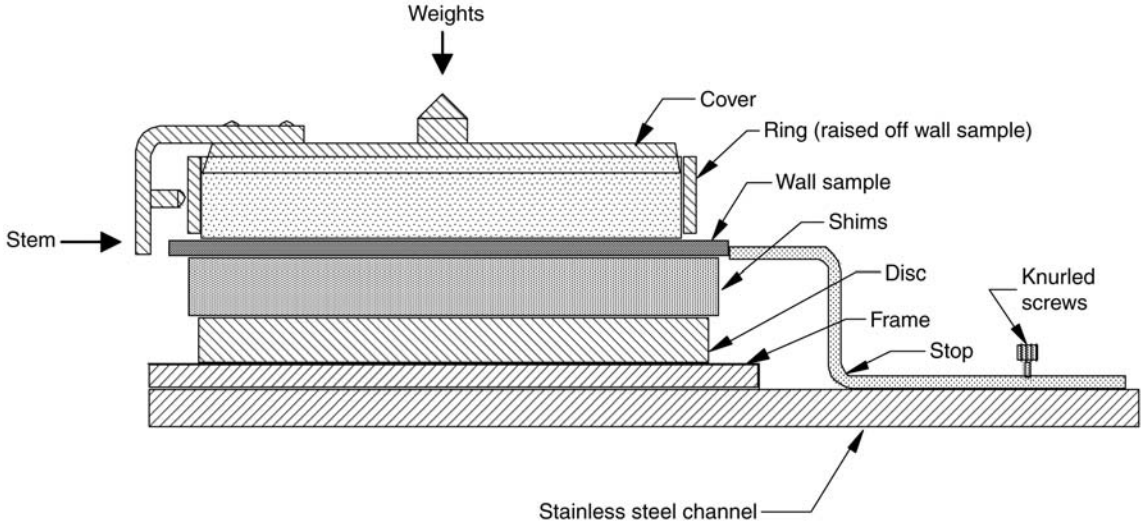


FIGURE 26.17 Jenike Direct Shear Test, wall friction test set-up.

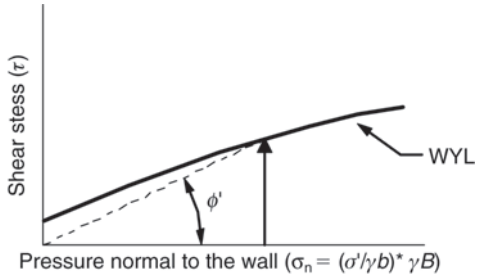


FIGURE 26.18 Example of wall yield locus generated from wall friction test data.

this case, a coupon of the wall material being evaluated is held in place on the frame of the tester, and a cell of powder is placed above. The coefficient of sliding friction (μ, μ = tangent ϕ') is the ratio of the shear force required for sliding (τ) to the normal force applied perpendicular to the wall material coupon (σ_n). A plot of the measured shear force (τ) as a function of the applied normal pressure (σ_n) generates a relationship known as the wall yield locus (see Fig. 26.18).

The wall friction measured is a function of the powder handled and the wall surface (type, finish, orientation) in contact with it. Variations in the powder, handling conditions (eg, temperature/RH), and/or the wall surface finish can have a dramatic effect on the resulting wall friction coefficient.[18] The results of the wall friction test are used to determine the hopper angles required to achieve mass flow and will be discussed in the next section.

26.4.1.5 Calculation of recommended mass flow hopper angles

Based upon mathematical models, design charts[9] have been developed to determine which flow pattern is to be expected during gravity discharge from a bin. The design charts use the following inputs for the powder being handled, and the bin design being considered:

1. the hopper angle ("θ_c" for a conical hopper or "θ_p" for a planar hopper), as measured from the vertical;
2. the wall friction angle (ϕ'), as measured from the wall friction tests;
3. the internal friction angle (δ), as measured from the cohesive strength tests.

This chapter will focus on the calculation of the recommended mass hopper angles for a conical hopper (θ_c) because the majority of pharmaceutical processes utilize bins with a conical hopper. The methods to calculate the recommended mass hopper angles for a planar hopper (θ_p) with a slotted outlet are

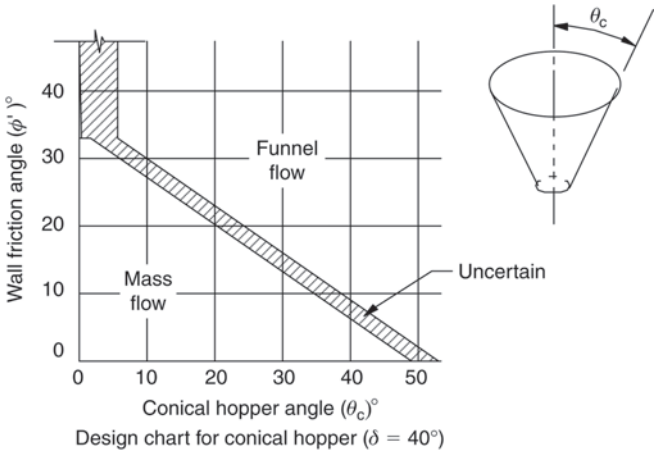


FIGURE 26.19 Mass flow/funnel flow design chart for conical hopper handling a bulk solid with an effective angle of internal friction (δ) of 40 degree.

similar in approach, and are outlined in Jenike.[9] It is worth noting that the recommended mass flow angles for planar hopper walls (θ_p) can often be 8–12 degree shallower than those for a conical hopper (θ_c), for the same sized opening.

An example of such a design chart for a conical hopper is shown in Fig. 26.19. The design chart shown is specifically for a powder with an effective angle of internal friction (δ) of 40 degree; the design charts will be different for different values of δ.[9] Hopper angles required for mass flow are a function of δ, since flow along converging hopper walls involves interparticle motion of the powder. For any combination of ϕ' and θ_c that lies in the mass flow region, mass flow is expected to occur. If the combination lies in the funnel flow region, funnel flow is expected. The "uncertain" region is an area where mass flow is expected to occur in theory, but represents a 4 degree margin of safety on the design, to account for inevitable variations in test results and surface finish.

As an example of using the design chart, a bin with a conical hopper angle (θ_c) of 20 degree from vertical is used. Wall friction tests are conducted on the hopper wall surface, and a wall friction angle of 35 degree is measured for the normal pressure calculated at the outlet. Based upon the design chart, this bin would be expected to discharge in funnel flow. In that case, the designer would need to find another wall surface, with a wall friction angle that is less than 20 degree, to ensure mass flow discharge.

The wall friction angle (ϕ') is determined by the wall friction tests described previously. The value of ϕ' to use for the hopper design charts will be selected for

the expected normal pressure (σ_n) against the surface at the location of interest in the bin (eg, the hopper outlet). For many combinations of wall surfaces and powders, the wall friction angle changes, depending on the normal pressure. When mass flow develops, the solids pressure normal to the wall surface is given by the following relationship:

$$\sigma_n = (\sigma'/\gamma b) * \gamma B \tag{26.4}$$

where:

$(\sigma'/\gamma b)$ is a dimensionless parameter that can be found in Jenike[9];
B (units of length) is the span of the outlet: the diameter of a circular outlet or the width of a slotted outlet;
is the bulk density at the outlet.

Generally, ϕ' increases with decreasing normal pressure (σ_n). The corresponding normal pressure to the wall (σ_n) is the lowest at the outlet where the span (B) is the smallest. Therefore, it is at the outlet where the wall friction angle (ϕ') is the highest for a given design, provided the hopper interior surface-finish and angle remain constant above the outlet. As a result, if the walls of the hopper are steep enough at the outlet to provide mass flow, mass flow should be expected at the walls above the outlet (regardless of total bin size).

The hopper angle required for mass flow is principally dependent on the outlet size selected for the hopper under consideration. The hopper angle required for mass flow is not a function of the flow rate, the level of powder within the bin or the diameter or height of the bin. A steeper hopper is often required to achieve mass flow for a bin with a smaller outlet because the wall friction angle generally increases with lower normal pressures.

26.4.2 Bulk density

The bulk density of a given powder is not a single or even a dual value, but varies as a function of the consolidating pressure applied to it. There are various methods used in industry to measure bulk density. One prominent method is utilizing different sized containers that are measured for volume after being loosely filled with a known mass of material ("loose" density), and after vibration or tapping (tapped density), such as the USP Chapter <616> method.[19] These methods can offer some repeatability with respect to the conditions under which measurements are taken. However, they do not necessarily represent the actual compaction behavior of a powder being handled in a bin, chute, or press hopper. Therefore, it is necessary to measure the bulk density over a range of consolidation pressures, via a compressibility test,[9,20] for design purposes. The results of the compressibility test can often be plotted as a straight line on a log–log plot (see Fig. 26.20). In powder handling literature, the slope of this line is typically called the "compressibility" of the powder.

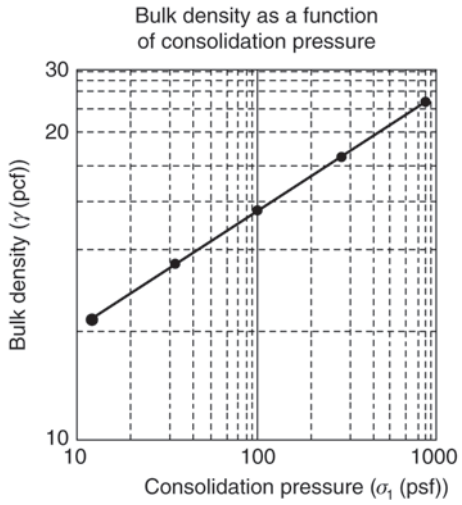


FIGURE 26.20 Example of bulk density versus consolidation pressure plot from compressibility test data.

The resulting data can be used to determine capacities for storage and transfer equipment, and evaluate wall friction and feeder operation requirements. As an example, when estimating the capacity of a bin, the bulk density based upon the average major consolidation pressure in the bin can be used. For the calculation of the arching dimensions (B_c), and recommended mass flow hopper angles (θ_c), the bulk density based on the major consolidation pressure at the bin outlet can be used.

26.4.3 Permeability

The flow problems that can occur due to adverse two-phase (powder and interstitial gas) flow effects were reviewed in Section 26.3. These problems are more likely to occur when the target feed rate (eg, tableting rate) exceeds the "critical flow rate," based on the powder's physical properties. The results of the permeability test are one of the primary flow properties used to determine the critical flow rate. The permeability of a powder is a measurement of how readily gas can pass through it. The permeability will have a controlling effect on the discharge rate that can be achieved from a bin/hopper with a given outlet size. Sizing the outlet of a

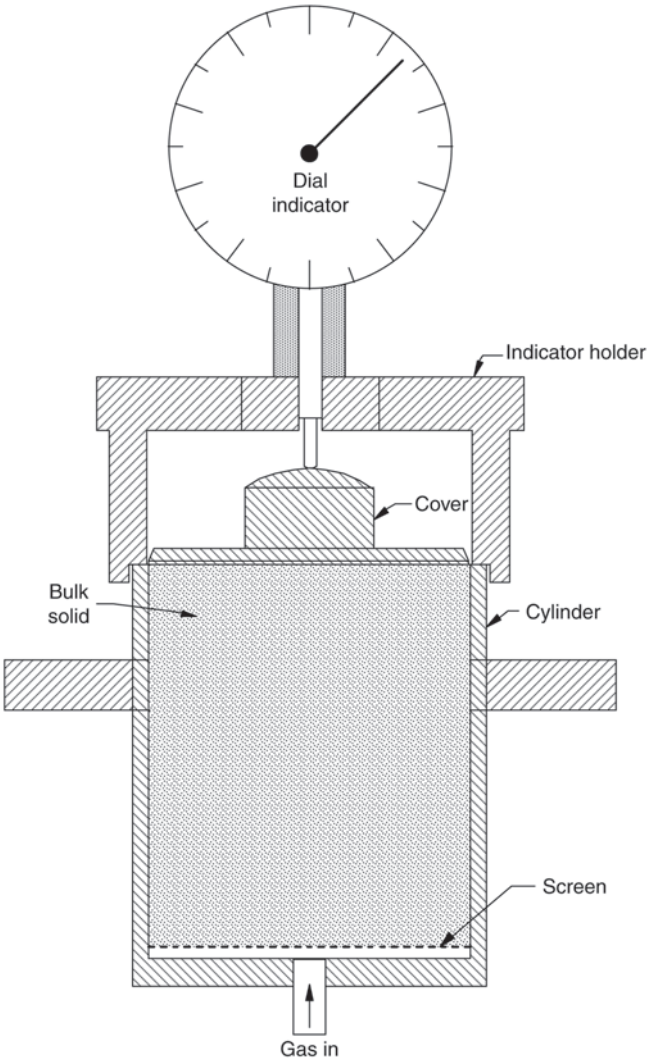


FIGURE 26.21 Schematic of permeability test set-up.

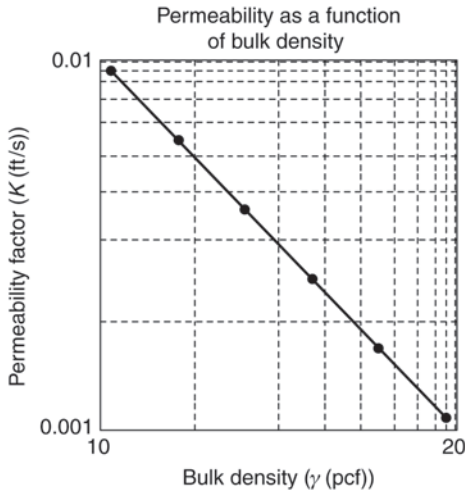


FIGURE 26.22 Example of permeability versus bulk density plot from permeability test data.

piece of equipment or choosing the diameter of a transfer chute should take into consideration the target feed rate.

Permeability is measured as a function of bulk density.[20] A schematic of the permeability test is provided in Fig. 26.21. In this test set-up, gas is injected at the bottom of the test cell through a permeable membrane. During the test, the pressure drop and flow rate across the powder are measured. The method involves measuring the flow rate of air at a predetermined pressure drop through a sample of known density and height. The permeability is then calculated using Darcy's law. The permeability of a powder typically decreases as the bulk density increases, so the test is conducted over a range of bulk densities.

Once the permeability/bulk density relationship is determined (see Fig. 26.22), it can be used to calculate the critical flow rates that will be achieved for steady-flow conditions though various outlet sizes. The critical flow rate is dependent upon the permeability, bin geometry, outlet size, and consolidation pressure. The details of calculating critical flow rates are outside the scope of this chapter, but mathematical models have been developed for these calculations.

Higher flow rates than the calculated critical flow rate may occur, but may result in unsteady or erratic feed and resulting adverse effects, as discussed in Section 26.3. Permeability values can also be used to calculate the time required for fine powders to settle or deaerate in equipment.

26.4.4 Segregation tests

When developing a product or designing a process, it is beneficial to know whether the blend will be prone to segregation. If the blend is prone to segregation, it is beneficial to know what segregation mechanism(s) will occur, since this information can be used to modify the material properties (eg, excipient selection, component PSD, etc.) to minimize the potential for segregation. An understanding of the potential for segregation can alert the equipment or process designer to potential risks that may then be avoided during the scale-up process. In some cases, significant process steps, such as granulation, may be required to avoid potential segregation problems.

There are two ASTM standard practices on segregation test methods.[21,22] These tests are designed to isolate specific segregation mechanisms, and test a material's tendency to segregate by that mechanism. A brief description of these test methods is provided next.

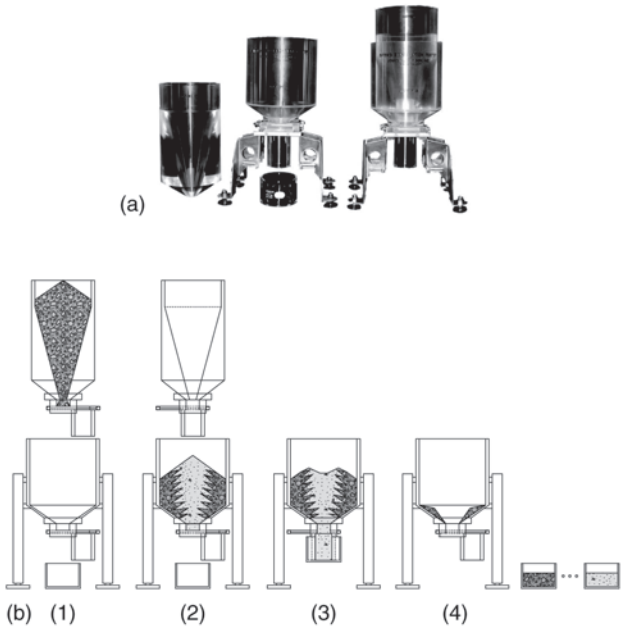


FIGURE 26.23 Sifting segregation testers (a); and sifting segregation test sequence (b). *Source: Photo provided courtesy of Jenike & Johanson, Inc.*

FIGURE 26.24 Fluidization segregation tester (controls not shown). *Source: Photo provided courtesy of Jenike & Johanson, Inc.*

26.4.4.1 Sifting segregation test method

The sifting segregation test (Fig. 26.23a) is performed by center-filling a small funnel flow bin, and then discharging it while collecting sequential samples. If sifting segregation occurs either during filling or discharge, the fines content of the discharging material will vary from beginning to end. Samples are collected from the various cups (ie, the beginning, middle, and end of the discharge). These collected samples can then be measured for segregation by particle size analyses, assays, and/or other variables of interest.

The sequence for performing the sifting segregation test is depicted in Fig. 26.23b, and is:

1. the blend is placed in mass flow bin;
2. the material is discharged from a fixed height, dropping into a funnel flow bin. This transfer of material will promote segregation if the material is prone to segregate due to sifting;
3. the material is discharged from the funnel flow bin. The discharge pattern will cause material from the center to discharge first, and material from near the walls to discharge last; and
4. the collected samples are then measured for segregation.

26.4.4.2 Fluidization segregation test method

The fluidization segregation test (Fig. 26.24) is run by first fluidizing a column of material by injecting air at its base. After the column is thoroughly fluidized, it is held near a minimum fluidization velocity for a predetermined period of time. The air is then turned off, and the material is allowed to deaerate. The column is then split into three equal sections (top, middle, and bottom) and the resulting samples are measured for segregation.

Several other researchers and companies have developed various segregation testers, including methods that induce vibration,[23,24] shearing in a cell,[25] and methods that capture material from a pile after it has formed.[25–27] Improvements and variations to the ASTM test methods have also been made. A fluidization segregation tester that utilizes a different mechanism to fluidize the bed has been developed.[28] It uses a smaller test sample, and provides unit-dose samples for analysis. As another example, an alternate way to run the ASTM sifting segregation test involves cycling the blend multiple times to strengthen the segregation "signal."[29]

Segregation tests are useful for identifying the:

1. segregation mechanism(s) that might be active for a given blend;
2. general segregation trend that may be observed in the process; and
3. comparisons between different formulas, variations of the same formula, etc.

However, the test results have limitations. Most notably, the segregation results are not scaleable from the tester to the manufacturing process, and therefore cannot be tied quantifiably to the process. The segregation test results do not necessarily mean that a highly segregating material cannot be handled in a manner that

prevents content uniformity problems. Therefore, the segregation test methods are primarily used as a stress test to identify the dominant segregation mechanism(s) expected to occur. This information enables the equipment designer to take the appropriate precautionary measures during scale-up or make corrective actions to existing equipment. Design techniques to minimize the potential for segregation are outlined in Section 26.5.

26.5 BASIC EQUIPMENT DESIGN TECHNIQUES

Objective of the section

The primary objective of Section 26.5 is to review basic design techniques for the bin-to-press feed system, to provide consistent, reliable flow for gravity feed, and minimize segregation. In particular, we will review the following basic design techniques:

1. reliable funnel flow bin design;
2. reliable mass flow designs for bins, transfer chutes, and press hoppers;
3. minimizing adverse two-phase flow effects (eg, feed/tableting rate limitations, flooding);
4. minimizing segregation during blender to press transfers.

For each of these different design concerns, we will review the key equipment design parameters. Regardless of whether the equipment being designed is a bin, transfer chute, or press hopper, a crucial first step in designing a reliable feed system is determining the flow pattern, and designing the equipment accordingly. These design techniques will focus on equipment commonly used in batch processes, but can also be applied to continuous manufacturing equipment. Section 26.4 discussed the wall friction tests and design charts used to determine if a hopper will discharge in mass flow or funnel flow.

26.5.1 Reliable funnel flow design (preventing a rathole)

Funnel flow occurs when the hopper walls are not smooth and/or steep enough to promote flow at the walls. Funnel flow bins can be prone to ratholing if the material is cohesive enough. A funnel flow bin design can be considered if all of the following design criteria are met:

1. Segregation of the final blend is not a concern. Since a funnel flow bin will discharge in a first-in-last-out flow sequence, any side-to-side segregation that occurred as the bin was filled will often be exacerbated in funnel flow discharge.
2. The final blend has low cohesive strength, therefore the formation of a stable rathole is not a concern. This can be checked by comparing the bin outlet diameter/diagonal length to the critical rathole diameter (D_f) for the estimated major consolidation pressure (σ_1) for the bin design in question as per Eqs. (26.2) and (26.3) (see Section 26.4). If the outlet diameter is less than D_f, ratholing is a concern.
3. Flooding due to a collapsing rathole is not a concern. Flooding can result in highly aerated (low density) powder being fed from the bin to the press, which may have an adverse effect on the tablet properties (weight, hardness, dissolution variation), and can result in segregation.
4. A nonuniform feed density is not a concern. Variation of the feed density into the press feed frame can result in tablet weight variation because tablet presses operate as volumetric feeders.

If all of these design criteria are met, a funnel flow bin design can be considered. If a funnel flow bin design is acceptable, the first concern is checking that the outlet diameter is greater than the critical rathole diameter (D_f), to ensure that a stable rathole will not form. If the diameter of the funnel flow bin is not greater than D_f, the following steps can be considered to reduce the likelihood of ratholing:

1. Enlarge the bin opening: this may require using a slotted outlet, which would require a feeder capable of feeding uniformly across the entire outlet (eg, mass flow screw feeder) or a valve capable of shutting off such an outlet. Using a larger outlet diameter may not be a practical modification, since the opening may need to be increased to be larger than standard valve or feeder sizes.
2. Reduce the material level in the bin: the critical rathole diameter typically decreases with a reduction in the major consolidation pressure (σ_1), which is a function of the fill height.
3. Use an internal, mechanical agitator: an internal, mechanical agitator, such as an agitator with "arms" that rotate about a central vertical shaft, may be a practical modification on a small scale for a press hopper. A bin with a discharge valve (eg, Matcon discharge valve) could also be considered a means of failing a stable rathole, but may need to be assessed via full scale trials to determine the operating parameters required (valve "stroke" setting, etc.).
4. Use external vibrators: the effectiveness of external vibrators to collapse a stable rathole would need to be assessed via full-scale trials prior to installation, since vibration may actually increase the strength on the blend in the bin and increase the likelihood of ratholing. Trials would be required to assess the optimum vibrator type

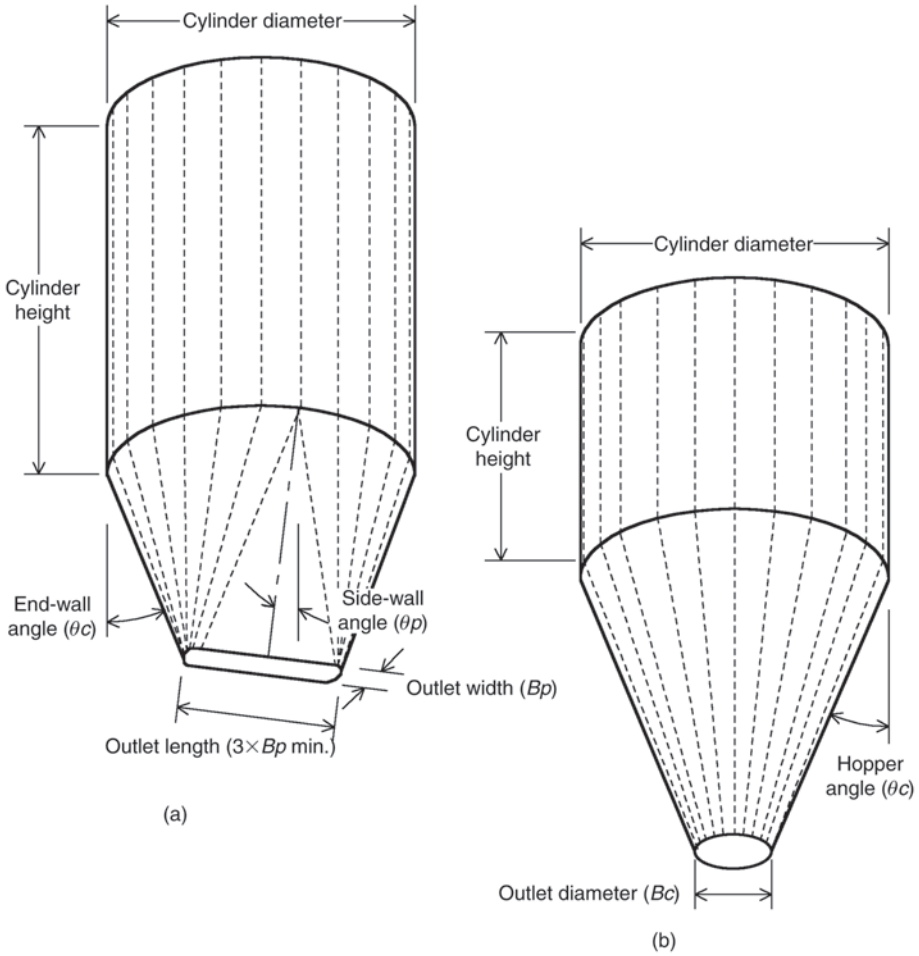


FIGURE 26.25 Mass flow design parameters (B_c, B_p, θ_c, θ_p). (a) Transition hopper; (b) Conical hopper

(high-frequency/low-amplitude vs low-frequency/high-amplitude), number of vibrators required, location, frequency settings, and other design considerations.

There are several adverse effects of using a bin that discharges in funnel flow (first-in-last-out flow sequence, nonuniform feed density, exacerbation of segregation). In addition, the potential options to prevent a rathole are often limited or impractical. Therefore, a common design technique for preventing ratholing is to redesign the bin for mass flow. The design techniques for mass flow are discussed in the following section.

26.5.2 Reliable mass flow designs for the bin, chute, and press hopper

Mass flow discharge from a bin occurs when the following two design criteria are met:

1. The bin walls are smooth and/or steep enough to promote flow at the walls;
2. The bin outlet is large enough to prevent an arch.

The wall friction tests and design charts used to determine if a bin will discharge in mass flow or funnel flow were discussed in Section 26.4. This section focuses on design techniques for mass flow bins, but these techniques may also be extended to obtaining mass flow in a transfer chute and press hopper. These techniques may be applied to designing new equipment or modifying existing equipment to provide mass flow.

When designing the bin to provide mass flow, the following general steps should be taken:

1. Size the outlet to prevent a cohesive arch: the bin designer should ensure that an arch will not form by making the outlet diameter equal to or larger than the minimum required outlet diameter (B_c, see Fig. 26.25). If a slotted outlet is used (maintaining a 3:1 length:width ratio for the outlet), the outlet width should be sized to be equal to or larger than the minimum required outlet width (B_p, see Fig. 26.25). The outlet may also need to be sized based upon the feed rate and two-phase flow considerations, as discussed next. If the outlet

cannot be sized to prevent an arch (eg, press hopper outlet that must mate with a fixed-feed frame inlet), an internal mechanical agitator or external vibrator could be considered, as discussed in the preceding section.

2. Make the hopper walls steep enough for mass flow: once the outlet is sized, the hopper wall slope should be designed to be equal to, or steeper than, the recommended hopper angle for the given outlet size, and selected wall surface. For a conical hopper, the walls should be equal to, or steeper than, the recommended mass flow angle for a conical hopper (θ_c, see Fig. 26.25). If the bin has a rectangular-to-round hopper, the valley angles should be sloped to be equal to or steeper than, θ_c. For planar walls, the walls should be equal to, or steeper than, the recommended mass flow angle for a planar hopper (θ_p, see Fig. 26.25).
3. Pay careful attention to the interior wall surface finish: when conducting the wall friction tests, it is beneficial to conduct tests on several different finishes (eg, #320 grit finish, #2B cold rolled finish, #2B electro-polished finish, etc.). This will provide the bin designer with a range of design options for the bin. Testing multiple wall surfaces will also enable the designer to assess the sensitivity of the wall friction results to different finishes. It is insufficient to simply test a 304 or 316 stainless steel with no regard to the interior finish. The wall friction of the blend may vary significantly from finish to finish. The orientation of directional finishes, such as a mechanical polish, is also critical to assess and control during fabrication. In addition, it cannot be assumed that an interior surface finish with a lower average roughness (R_a) will provide the best wall friction properties.
4. Consider velocity gradients: even when a bin is designed for mass flow, there still may be a velocity gradient between the material discharging at the hopper walls (moving slower) versus the center of the hopper (moving faster), assuming a symmetric bin with a single outlet in the center. Depending upon the application, the bin designer may want to increase the velocity gradient to enhance blending between vertical layers of material in the bin. Alternatively, the bin designer may want to reduce the velocity gradient to enhance blending on a side-to-side basis. The decision to increase or decrease the velocity gradient will be depend upon the segregation that occurs upon filling the bin, and its effect upon content uniformity. The velocity gradient is reduced by making the hopper slope steeper, with respect to the recommended mass flow hopper angle (θ_c). The velocity gradient is increased by making the hopper slope closer to (but still steeper than) the recommended mass flow hopper angle. Changing the interior surface to reduce friction or using an insert (discussed more next) are other methods used to control the velocity gradient. Asymmetric hoppers, which are common for press hoppers, are especially prone to velocity gradients since the material will move faster at the steeper hopper wall. In addition, a velocity gradient cannot be completely eliminated, especially as the material level in the hopper empties. Velocity profiles, and their effect on blending material, can be calculated a priori given the geometry of the bin (θ_c) and measured flow properties that were discussed in Section 26.4 (ie, ϕ', δ, ϕ).
5. Avoid upward-facing lips/ledges due to mismatched flanges (eg, see Fig. 26.26), level probes, view ports, partially opened valves, etc., especially in the hopper section. Interior protruding devices should be ideally located in the straight-sided (nonconverging) section of a bin/press hopper if possible, where they will be less detrimental in upsetting a mass flow pattern.

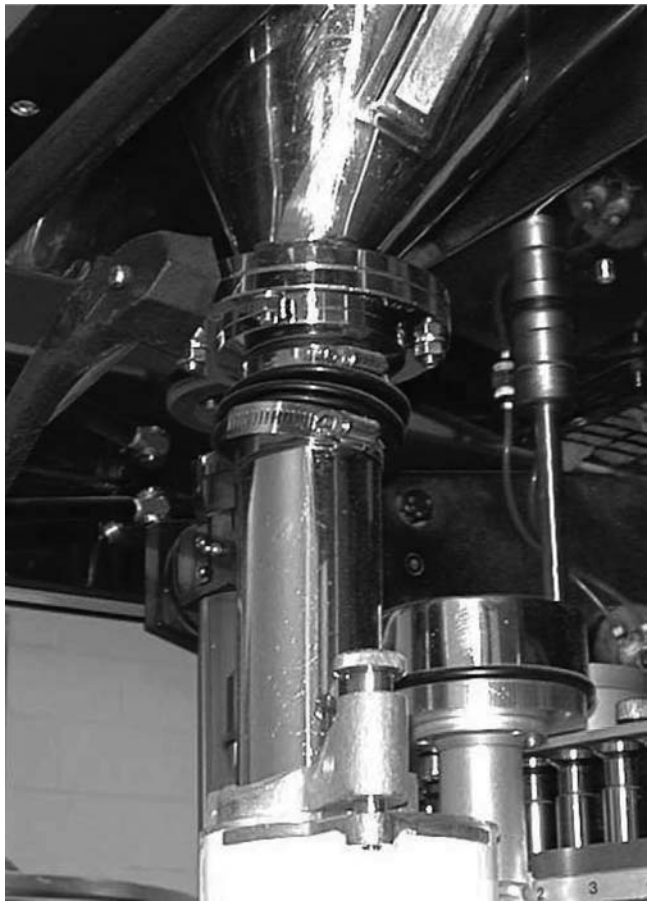
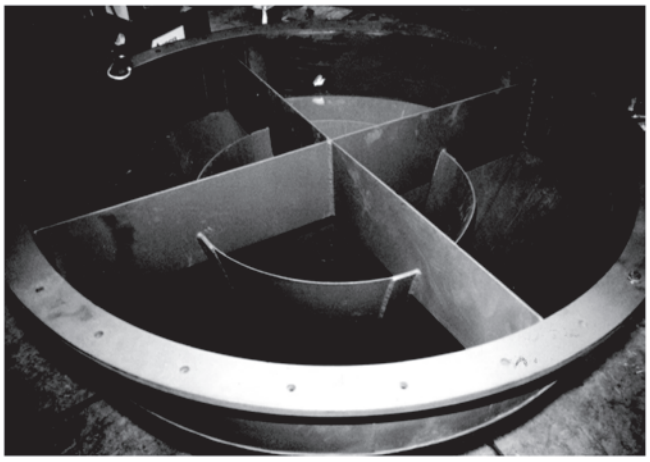


FIGURE 26.26 Example of an upward-facing ledge at a flange connection.

If the bin designer is modifying an existing funnel flow bin to provide mass flow, several different options can be considered, including:

1. Using a different interior surface finish with better wall friction properties (ie, lower wall friction). The bin designer should conduct wall friction tests on alternative wall surfaces, to assess if changing the surface finish (eg, electro-polishing an existing #2B finish) will convert the bin from funnel flow to mass flow. This is often one of the most cost-effective modifications to obtain mass flow.
2. Using a flow-controlling insert such as a Binsert (Fig. 26.27) to obtain mass flow within the same bin. A properly designed insert can change the stresses that develop in the bin during discharge, so that mass flow can be obtained at a wall where the material was previously stagnant.
3. Modifying the hopper geometry. Use a different geometry that is more likely to provide mass flow (eg, conical instead of a rectangular-to-round hopper with shallower valley angles). If the hopper is modified to have a slotted outlet, it is crucial that the feeder the hopper mates to withdraws material across the entire outlet.



(a)

(b)

FIGURE 26.27 (a) Example of an "open" Binsert design; (b) Example of a "closed" Binsert design.

In addition to these design techniques for bins, there are several additional design techniques for designing transfer chutes for reliable mass flow including:

1. For converging sections that have a full cross-section (ie, hoppers) use the same design criteria discussed previously for a mass flow bin.
2. For nonconverging sections of the chute, the chute should be sloped to exceed the wall friction angle (ϕ') by a least a 10 degree margin of safety. As an example, if the measured wall friction angle for the given wall surface (from the wall friction test results) is 40 degree from horizontal, the recommended chute angle for the nonconverging portion of the chute would be at least 50 degree from horizontal.
3. If a bifurcated chute is used, the sloping chute legs should be symmetric, to prevent velocity gradients and the possibility of stagnant material in the shallower leg.
4. Use mitered joints between sloping and vertical sections.

26.5.3 Minimizing adverse two-phase flow effects

The primary focus in preventing adverse two-phase flow affects is to ensure that the powder handling equipment is designed so that the critical flow rate through a given outlet is greater than the target feed rate. The critical flow rate is determined using mathematical models, with permeability and compressibility test results as primary inputs. The critical flow rate is also a strong function of the outlet size, and increases as the outlet size increases. The target feed rate is often set by the required maximum tableting rate for the process. Adverse two-phase flow effects are typically most pronounced at the press feed hopper because it often has the smallest outlet in the entire press feed system. Therefore, the press hopper will typically have the lowest critical flow rate.

When designing the powder handling equipment to minimize adverse two-phase flow effects, the following general design techniques are beneficial:

1. Design the equipment for mass flow: mass flow will provide consistent feed and a more uniform consolidation pressure acting on the powders. In

addition, having a first-in-first-out flow sequence will allow the material more time to deaerate before being discharged through the outlet. This will reduce the likelihood of flooding. Mass flow will also prevent collapsing ratholes that can result in the powder aerating and flooding as it falls into the central flow channel. It is worth noting that mass flow can result in a lower critical flow rate than funnel flow, but will be more stable. Therefore, simply using a mass-flow bin design may not be the only corrective action required if a flow rate limitation occurs. However, designing the equipment for mass flow is often the first step in addressing adverse two-phase flow effects.

2. Use larger outlets for the handling equipment: the critical flow rate is a strong function of the cross-sectional area of the outlet. Therefore, increasing the outlet can often be highly beneficial in reducing two-phase flow effects. The goal would be to increase the outlet size until the critical flow rate for the selected outlet size exceeds the target flow rate. Since this may not be feasible for a press feeder hopper, in which the outlet size is fixed to mate with the press feed frame inlet, additional design techniques are discussed next. Computer software can be used to model the two-phase flow behavior to assess the effect of changing the outlet diameter.
3. Reduce the fill height in the handling equipment: the critical flow rate through a given outlet increases as the major consolidation pressure (σ) decreases. Therefore, reducing the fill height will be beneficial, but much less effective than increasing the outlet size.
4. Reduce the target feed rate: if possible, reducing the target feed rate (tableting rate) to be less than the critical flow rate will be beneficial, but is often impractical because it will result in decreased production rate.
5. Consider gas pressure differentials: a gas pressure differential can have a beneficial or adverse effect upon two-phase flow effects. A positive gas pressure differential at the outlet (ie, bin at a higher gas pressure than the equipment downstream) may be beneficial in overcoming a feed rate limitation. In this case, the air pressure is forcing the material in the direction of flow. Conversely, a negative gas pressure differential at the outlet can further reduce the critical flow rate because the negative gas pressure acts to further retard the flow rate.
6. Add air permeation: air permeation may be added to the system actively via an air injection system or passively through a vent. In particular, adding judicious (often very small) amounts of air at the location in the press feed system where the interstitial gas pressure is lowest can often be beneficial. However, this can be very unstable for small systems and/or very low permeability materials.
7. Changing the PSD of the powder: the permeability of a powder is a strong function of its PSD. Powders with a finer PSD are often less permeable and, therefore, more prone to adverse-two phase flow effects. Even a reduction in the percentage of fine PSDs can often be beneficial in increasing the permeability of a powder, and decreasing the likelihood of adverse two phase-flow effects.

The key to implementing any corrective actions designed to reduce adverse flow-effects will be using a mathematical two-phase flow analysis to assess the effects on the bulk solids' stresses and interstitial gas pressure. This analysis would need to use inputs such as key flow properties (ie, permeability, compressibility), and equipment/process parameters (ie, tableting rate, bin/hopper geometry, and gas pressure gradients) to assess the effect of the potential corrective actions outlined previously.

26.5.4 Minimizing segregation in the blender-to-press transfer steps

It may be challenging for a designer to determine which segregation mechanism(s) is dominant, and develop appropriate corrective actions. This requires knowledge of the material's physical and chemical characteristics, as well as an understanding of the segregation mechanisms that can be active. One must identify the process conditions that can serve as a driving force to cause segregation. Flow properties measurements (ie, wall friction, cohesive strength, compressibility, and permeability) can help to provide understanding of the behavior of the material in storage and transfer equipment. Consideration should be given to the fill/discharge sequence, including flow pattern and inventory management, which gives rise to the observed segregation. Testing for segregation potential (see Section 26.4) can provide additional insight about the mechanisms that may cause segregation. Sufficient sampling is required to support the hypothesis of segregation (eg, blend samples and final product samples, samples from the center vs periphery of the bin). Finally, one must consider the impact of analytical and sampling errors specific to the blend under consideration, as well as the statistical significance of the results, when drawing conclusions from the data.

From the previous discussion about segregation mechanisms, it can be concluded that certain material properties, as well as process conditions, must exist for segregation to occur. Elimination of one of these will prevent segregation. It stands to reason, then, that if segregation is a problem in a process, one should look for opportunities to either: (1) change the material or

(2) modify the process equipment or conditions. This chapter focuses on the equipment and process design techniques to minimize segregation in batch processes.

Some generalizations can be made when designing equipment to minimize segregation. The complete details on how to implement these changes correctly are beyond the scope of this chapter. However, all equipment must be designed based on the flow properties and segregation potential of the blends being handled.

Primary equipment and process design techniques to minimize segregation during the final blender-to-press transfer include:

1. Minimize the number of material transfer steps. The tendency for segregation increases with each transfer step and movement of the bin. Ideally, the material would discharge directly from the blender into the tablet press feed frame, with no additional handling. In-bin blending is as close to this as most firms can obtain practically. This assumes that a uniform blend can be obtained within the bin blender in the first place.
2. Storage bins, press hoppers, and chutes should be designed for mass flow. In mass flow, the entire contents of the bin are in motion during discharge. In funnel flow, stagnant regions exist.
3. Minimize transfer chute volumes to reduce the volume of displaced air, and the volume of potentially segregated material. However, the chute must remain large enough to provide the required throughput rates.
4. Use a storage bin with a larger height:diameter aspect ratio. A mass flow bin with a tall, narrow cylinder will minimize the potential for sifting segregation, compared with a short, wide bin. A disadvantage is that the taller drop height may exacerbate other segregation mechanisms.
5. Bins and blenders should be vented to avoid gas counterflow. Air in an otherwise "empty" bin must be displaced out of the bin as powder fills it. If this air is forced through material in the V-blender, it can induce fluidization segregation within the blender. A separate pathway or vent line to allow the air to escape without moving through the bed of material can reduce segregation.
6. Velocity gradients within bins should also be minimized. The hopper must be significantly steeper than the mass flow limit to achieve this. A steeper hopper section may result in an impractically tall bin. Alternate approaches include the use of inserts (discussed previously). If an insert is used, it must be properly designed and positioned to be effective. Asymmetric bins and hoppers should be avoided if possible, and symmetrical ones should be used whenever possible.
7. Dust generation and fluidization of the material should be minimized during material movement. Dust can be controlled by way of socks or sleeves, to contain the material as it drops from the blender to the bin, for example. Some devices are commercially available. An example of this is a solids decelerator shown in Fig. 26.28.
8. Drop heights should be minimized where possible. Drop heights may aerate the material, induce dust, and increase momentum of the material as it hits the pile. This will increase the tendency for each of the three segregation mechanisms to occur.
9. Valves should be operated correctly. Butterfly valves should be operated in the full open position, not throttled to restrict flow. Restricting flow will virtually assure a funnel flow pattern, which is usually detrimental to uniformity.
10. Use a symmetrical split whenever a process stream is divided. A symmetrical split, such as a bifurcated chute to feed two sides of a press (a Y-branch), will eliminate potential differences in the flow between the two streams. Consideration must be given to any potential for segregation upstream of the split. Even seemingly minor details, such as the orientation of a butterfly valve prior to a split, can affect segregation. Proper designs should be utilized for Y-branches to avoid stagnant material and air counterflow.

FIGURE 26.28 Example of a solids deceleration device. *Source: Photo provided courtesy of GEA Process Engineering.*

Other specific solutions may become apparent once the segregation mechanism has been identified. For example, mass flow is usually beneficial when handling segregation-prone materials, especially materials that exhibit a side-to-side (or center-to-periphery) segregation pattern. Sifting and dusting segregation mechanisms fit this description.

It is important to remember that mass flow is not a universal solution because it will not address a top-to-bottom segregation pattern. As an example, consider the situation in a portable bin where fluidization upon filling the bin has caused the fine fraction of a blend to be driven to the top surface. Mass flow discharge of this bin would effectively transfer this segregated material to the downstream process, delivering the coarser blend first, followed by the fines.

In summary, when addressing segregation concerns it is crucial to know the chosen process, and how the blend will segregate before implementing equipment designs or corrective actions.

References

1. Rippie EG. Powders. In: Osol A, Chase GD, Gennaro AR, Gibson MR, Granberg CB, editors. *Remington's pharmaceutical sciences*. 16th ed. Easton, PA: Mack Publishing Company; 1980. . p. 1535–52.
2. Venables HJ, Wells JI. Powder mixing. *Drug Dev Ind Pharm* 2001;**27**(7):599–612.
3. Alexander AW, Muzzio FJ. *Batch size increase in dry blending and mixing*. 2nd edn. *Pharmaceutical process scale-up*, 2006. New York, NY: Taylor & Francis; 2006. p. 161–80.
4. Royal TA, Carson JW. Fine powder flow phenomena in bins, hoppers and processing vessels. Presented at Bulk 2000. London; 2000.
5. Bates L. *User guide to segregation*. Marlow: British Materials Handling Board; 1997.
6. Williams JC. The segregation of particulate materials: a review. *Powder Technol* 1976;**15**:245–51.
7. Williams JC, Khan MI. The mixing and segregation of particulate solids of different particle size. *Chem Eng* 1973;**19**:19–25.
8. Pittenger BH, Purutyan H, Barnum RA. Reducing/eliminating segregation problems in powdered metal processing. Part I: segregation mechanisms. *P/M Sci Technol Briefs* 2000;**2**(1):5–9.
9. Jenike AW. *Storage and flow of solids. Bulletin 123 of the Utah engineering experimental station*, vol. 53(26). Salt Lake City, UT: University of Utah; 1964, Revised 1980
10. Schulze D. Measuring powder flowability: a comparison of test methods part I. *Powder Bulk Eng* 1996;**10**(4):45–61.
11. Schulze D. Measuring powder flowability: a comparison of test methods part II. *Powder Bulk Eng* 1996;**10**(6):17–28.
12. Standard shear testing method for powders using the Jenike shear cell. ASTM Standard D6128-06, American Society for Testing and Materials; 2006.
13. Bausch A, Hausmann R, Bongartz C, Zinn T. Measurement of flowability with a ring shear cell, evaluation and adaptation of the method for use in pharmaceutical technology. In: *Proceedings of the 2nd world meeting APGI/APV*, Paris, May 1998. p. 135–6.
14. Hausmann R, Bausch A, Bongartz C, Zinn T. Pharmaceutical applications of a new ring shear tester for flowability measurement of granules and powders. In: *Proceedings of the 2nd world meeting APGI/APV*, Paris, May 1998. p.137–8.
15. Nyquist H. Measurement of flow properties in large scale tablet production. *Int J Pharm Technol Product Manuf* 1984;**5**(3):21–4.
16. Ramachandruni H, Hoag S. Application of a modified annular shear cell measuring lubrication of pharmaceutical powders. University of Maryland, Poster presentation at the AAPS annual meeting, San Francisco; 1998.
17. Janssen HA. Versuche uber Getreidedruck in Silozellen. *Verein Deutcher Igenieure, Zeitschrift* 1895;**39**:1045–9.
18. Prescott JK, Ploof DA, Carson JW. Developing a better understanding of wall friction. *Powder Handling Process* 1999;**11**(1):27–35.
19. The United States Pharmacopeia. Chapter, 616. Bulk and tapped density. *US Pharmacopeial Forum* 2002;**28**(3).
20. Carson JW, Marinelli J. Characterize powders to ensure smooth flow. *Chem Eng* 1994;**101**(4):78–90.
21. Anonymous. (2003). Standard practice/guide for measuring sifting segregation tendencies of bulk solids. D6940-03, ASTM International.
22. Anonymous. (2003). standard practice for measuring fluidization segregation tendencies of powders. D6941-03, ASTM International.
23. Ahmed H, Shah N. Formulation of low dose medicines—theory and practice. *Am Pharm Rev* 2000;**3**(3):9–15.
24. Globepharma's Powdertest ™ <http://www.globepharma.com/html/powertest.html> [accessed September 2006].
25. Johanson K, Eckert C, Ghose D, et al. Quantitative measurement of particle segregation mechanisms. *Powder Technol* 2005;**159**(1):1–12.
26. Massol-Chaudeur S, Berthiaux H, Doggs J. The development and use of a static segregation test to evaluate the robustness of various types of powder mixtures. *Trans IchemE* 2003;**81** (Part c):106–18.
27. De Silva SR, Dyroy A, Enstad GG. Segregation mechanisms and their quantification using segregation testers. In: Solids mechanics and its applications, volume 81, IUTAM symposium on segregation in granular flows; 1999. p. 11–29.
28. Hedden DA, Brone D, Clement SA, et al. Development of an improved fluidization segregation tester for use with pharmaceutical powders. *Pharm Technol* 2006;**30**(12):54–64.
29. Alexander B, Roddy M, Brone D, et al. A method to quantitatively describe powder segregation during discharge from vessels. *Pharm Technol Yearbook* 2000;**2**:6–21.

CHAPTER

27

Capsules Dosage Form: Formulation and Manufacturing Considerations

S.W. Hoag

University of Maryland, Baltimore, MD, United States

27.1 INTRODUCTION—CAPSULES AS A DOSAGE FORM

Capsules are a unique dosage form with a long history of use in pharmacy. The original patent was issued in 1834 to a Parisian pharmacist, Joseph Gérard Dublanc, and pharmacy student, François Achille Barnabé Mothès, for the invention and manufacture of gelatin capsules.[1] The basic idea of a capsule is to enclose the drug or active pharmaceutical ingredient (API) in an odorless, tasteless, elegant, easy-to-swallow, and easy-to-fill shell. Today there are two main types of capsules: the hard gelatin capsule and the soft gelatin capsule, often called softshells. This chapter will focus on hard gelatin capsules. The hard gelatin capsule can be used for dry fills such as powder, liquids, and semisolids, while the softshell is exclusively used for liquids and semisolids. The typically capsule shell is made of gelatin, but in recent years, there have been a variety of gelatin alternatives introduced to the market. The vast majority of capsule applications are for oral delivery of an API; however, there are specialty applications such as capsules that can be loaded into dry-powdered inhalers, add reagents as part of a diagnostic kit, and occasionally as a suppository base with glycerin.[2] The majority of capsules are filled with a dry powder; however, semisolids, nonaqueous liquids, and other dosage forms such as beads, mini tablets, and even mini capsules can be filled into a capsule shell. These applications will be discussed. In terms of production, capsules are one of the most flexible forms. They can be made one a time in a compounding pharmacy, in small-scale production for clinical studies, and all the way up to commercial production with machines that can make hundreds of thousands of capsules per hour. Given the importance capsules have in the pharmaceutical industry, the objective of this chapter is to give a broad overview of capsules, gelatin, capsule manufacture, capsule filling, and capsule formulation.

From the patient perspective, capsules have many advantages, making them among the most popular dosage forms on the market. Generally speaking, most patients consider capsule shells to be smooth, slippery, and easier to swallow than tablets.[3] In addition, capsules eliminate all contact between the drug and the mouth, which makes them tasteless and odorless, as most drugs have a bitter taste with an unpleasant aftertaste. This greatly improves patient compliance and consequently therapeutic outcomes. In addition, capsules can be made with a clear, high-gloss film that can be attractively colored and printed on, and they can have an elegant appearance, which also enhances patient acceptance. For patients who have trouble swallowing, the capsule can be opened up, and the contents sprinkled on a food such as applesauce, which is advantageous for pediatric and geriatric patients.

From a drug delivery point of view, capsules have many advantages. For immediate-release (IR) dosage forms, a key step is the breakdown of the capsule shell, which is analogous to disintegration in a tablet. For capsules, this occurs readily; see the discussion of gelatin cross-link in Section 27.2.2. Thus, capsules are ideally suited for IR delivery. In addition, capsules can be used for other types of release profiles as well. In addition to dry powder fills, multiparticulate beads can be filled into capsules, as shown in Fig. 27.1. For example, morphine sulfate has a short half-life, and with an IR delivery system, it requires dosing every 8 hours, but with a

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00027-3

FIGURE 27.1 The different types of fills that can be put into hard gelatin capsule shells. *Source: http://www.epmmagazine.com/news/empty-hard-capsules-and-barrier-packaging-films/.*

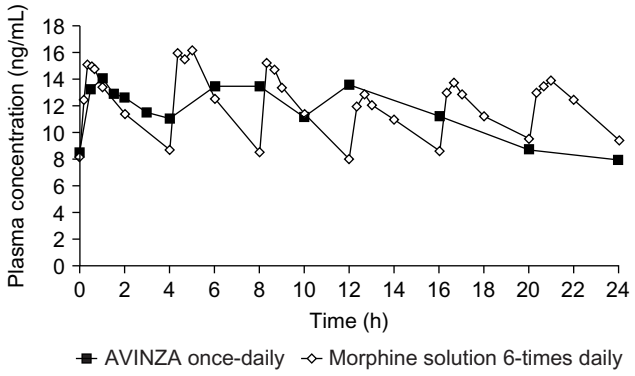


FIGURE 27.2 Multiparticulate dosing of morphine sulfate with Avinza, from the official FDA label.

controlled-release dosage such as Avinza or Kadian, it can be dosed once a day, as seen in Fig. 27.2, which is a big advantage in terms of compliance. For patients in an institutional environment, it has greater convenience for the nursing staff. In addition, with multiparticulate beads, you can have a mixture of beads with different release rates. If you look at the initial phase of the plasma concentration versus time pharmacokinetic profile in Fig. 27.2, you can see the initial onset is the same as the oral solution. This is because the coated beads are mixed with uncoated beads for rapid onset of pain relief. In addition to putting multiparticulate beads in a capsule, with modern filling equipment, you can also fill capsules with other dosage forms such as mini tablets and all possible combination of beads, tablets, capsules, powders, and even liquids, as shown in Fig. 27.1.[4]

From the company perspective, capsules can be used throughout the entire drug development process from preclinical studies in rodents to first-in-human studies all the way through the commercial manufacturing of a marketed product. While capsules are not as cost effective as tablets to manufacture, capsules are still relatively inexpensive to manufacture, and being a solid dosage form with a dry powder fill, capsule formulations promote good stability compared to other delivery systems such as liquid formulations. Another advantage of capsules is the relative ease of formulation. For initial formulations, the formulation requirements are minimal; however, as the speed of production goes up, the formulation requirements also go up. For capsules, the formulation requirements are that a consistent plug with a consistent weight is formed. For tablets, the formulation must flow well and be very compressible, which is a more stringent requirement than forming a plug with a consistent weight. The relative simplicity of capsule formulations can speed up the development process, especially for high-dose, poorly compressible drugs. This formulation simplicity makes capsules popular for clinical trials where speed is critical. Perhaps one of the biggest advantages of capsules is that they can be self-administrated, which is required for a drug to be a top seller, because of the high cost of administration by a health care worker such as with parenteral products. Nonself-administration products that are big sellers are typically only used for life-threating diseases such as diabetes where the patient must regularly give himself or herself an injection.

One big advantage of capsules from the marketing and patient perspective is the wide range of colors, sizes, and printings that are available for capsules shells (see Fig. 27.3). Having unique colors or printings of a logo can help companies establish the brand ID of a product for marketing purposes. A good example of this is Nexium (esomeprazole magnesium), which is an over-the-counter proton-pump inhibitor for gastroesophageal reflux. The unique "purple pill" has been successfully used to establish a recognizable product brand, and this could not be done without the unique purple color of the capsule shell and the printing of the gold bands and name on the capsule shell (see Fig. 27.3). Also, color is the most remembered trait of a medication and important for patient acceptance. For example, you wouldn't want to make an antidepressant capsule black or a sleep aid sunrise yellow and red. This is one area that capsules can easily accommodate a wide range of consumer tastes and preference requirements to produce an elegant dosage form that is appealing to patients.

Despite many advantages, capsules also have some important disadvantages. In common with all oral delivery systems, the capsule exposes the drug to the gastrointestinal (GI) tract and the GI tract to the drug. For example, some drugs when exposed to the GI tract

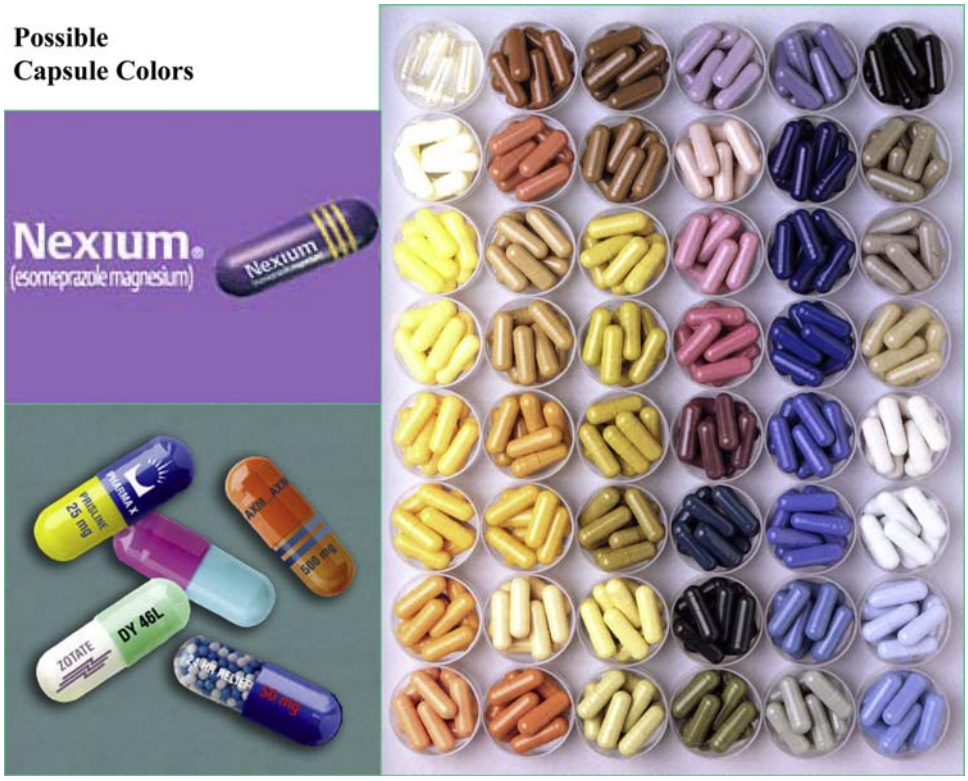


FIGURE 27.3 Diversity of capsule shell showing printing and colors.

can cause nausea, vomiting, and diarrhea. Other drugs such as iron in prenatal supplements and fish oil capsules have an unpleasant "burp back" effect, which reduces patient compliance. In addition, some drugs that are readily degraded in the GI tract, either due to stomach acids or enzymatic degradation in the stomach and small intestine, don't make good candidates for capsules. Also, for controlled-release dosage forms, the maximum duration of action is limited to the GI transit time.

Some disadvantages unique to capsules include interactions between the drug, the excipients, and the capsule shell. Hydroscopic materials may dry out the capsule shell and cause it to become brittle. A typical capsule shell has between 12% and 16% water, and if the water content drops much lower, the shell can crack during normal handling. Conversely, the capsule shells can absorb water from the environment and develop problems with drug stability due to too much water, and the capsule shell can become tacky. Another disadvantage of gelatin capsule shells is you have to make sure the shells are made from a bovine spongiform encephalopathy (BSE)–free source. In addition, capsule plugs are made with a lot less force than tablets, so for high-dose bulky materials such as botanical supplements, the dosage form size will be much bigger, and patients prefer smaller dosage forms. Tablets are less expensive to manufacture because you don't have to buy the capsule shells, and tablets have higher rates of production, which reduces capital expenditures. In my experience, capsule-filling machines are more complex than a tablet press and take a longer time to setup up, break down, and do cleaning validation due to all the moving parts. When doing a change over from one capsule size to another, you need to replace the dosing disk, the tamping pins or dosator, and all the bushings involved with separating the capsule shell. The change over from one size to another can take many hours to complete.

27.2 GELATIN AND CAPSULE SHELL COMPOSITION

Today, gelatin is the most commonly used material to make hard capsule shells; however, as gelatin alternatives gain in popularity, this may change in the future. Since the first patent in 1834, gelatin has been used to make capsule shells, and it has a long history of use in pharmaceutical capsules and as a food additive. Gelatin is cost effective and nontoxic with an excellent safety profile; the issue of BSE is discussed in Section 27.2.4. In addition, gelatin has many excellent mechanical properties such as the ability to form films

with precise dimensions and mechanical properties that can withstand the rigors of high-speed processing on commercial-scale production equipment. With gelatin, the film formation and gelling properties occur over a narrow temperature range.[5] The tribology surface properties and shell clarity are also critical for patient acceptance of a capsule shell. Also, gelatin has many excellent delivery properties. It is readily soluble in water and softens at body temperatures for rapid release; the issue of gelatin cross-linking is discussed in Section 27.2.2.[6] Gelatin is recognized in all the major pharmacopeias throughout the world. For these reasons, gelatin is the most popular material used to make two-piece hard capsule shells.

Gelatin is made up of proteins derived from the skin and the bones of cattle and pigs and is a byproduct of the meat industry, which helps make it cost effective. Gelatin is made up of primarily collagen, which is one of the main proteins in connective tissue. For example, 46% of pig skin and 29.4% of bovine hides are made of collagen.[6] Gelatin undergoes reversible thermal gelation at 35°C. This temperature and a narrow transition range are critical because capsules are made by a dipping process in a gelatin solution and when cooled must rapidly form a thin, uniform film on the dipping pin.[7] In addition, gelatin is readily soluble in water and softens at 30°C. These properties make it an excellent ingredient from a manufacturing and drug delivery point of view.[6]

There are different types of collagen found in different parts of the body. For example, skin is mainly made up of type I collagen and to a lesser extent type III collagen.[6] Collagen is composed of three chains that form a triple-helix structure. The chains have a high percentage of the amino acids glycine, proline, and hydroxyproline in a repetitive sequence.[6] When the three chains are linked together, they form a very stable triple helix that is stabilized by intrachain and interchain hydrogen bonds (see Fig. 27.4). Upon heat treatment, the triple helix can denature (other factors can lead to denaturation, but they are not relevant for this discussion), and depending upon temperature and water content, the chains can exist in different states with different degrees of cross-linking via the formation of a triple helix (see Fig. 27.5). Once denatured, they can form gels, and the extent of gel formation depends upon the temperature, moisture content, and how fast the gel solution was dried.[6,8,9] It is these changes in state with different degrees of cross-linking/triple-helix formation that contribute to the gelling behavior of gelatin and give gelatin its unique properties for forming capsules and drug release.

Gelatin is made by either acid or alkaline processing of cow bones or cow and pig hides. The process has many steps and requires a long time to complete.

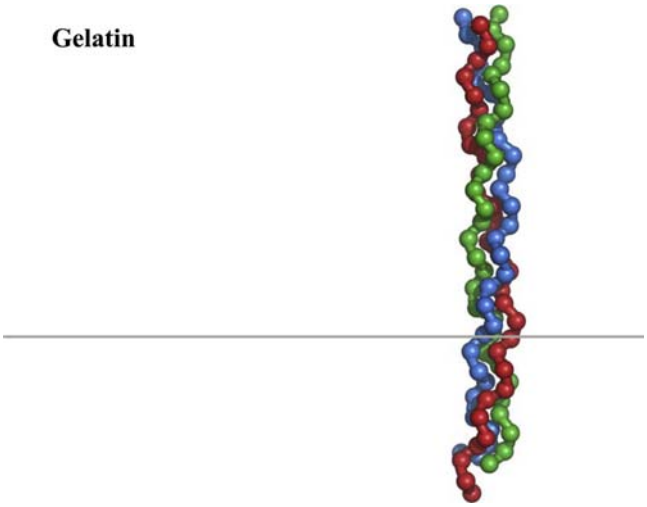


FIGURE 27.4 Structure of collagen. *Source: Wikipedia collagen.*

Fig. 27.6 shows the basic steps.[10] The gelatin made from pork skin via acid treatment is called Type A gelatin, and the gelatin made from bones and animal skin by alkaline processing is called Type B gelatin. Each gelatin type has its own properties (see Table 27.1). Type A gelatin has a higher degree of plasticity and clarity, and Type B gelatin forms tougher films and is a little hazier. A capsule shell is made from the mixture of the two gelatins, so the capsule shell has the right clarity and mechanical strength properties to run well on a filling machine.

The first step in the gelatin-making process is to pretreat the bones and hides (see Fig. 27.6). For bones, this means crushing, degreasing, and demineralization of the bones to remove all the calcium phosphate, $Ca_3(PO_4)_2$, using diluted hydrochloric acid (HCl) to solubilize the tricalcium phosphate, which can then be removed by washing. The spongy organic material left after the demineralization of the bones is called ossein.[11] The goal of the pretreatments is to produce a relatively purified collagen than can be converted into gelatin. After pretreatment, the gel can either be limed (alkaline treatment) or acid treated. The goal of liming is to condition the collagen so it produces gelatin with the desired physical properties and to do so with good yields.[10] The liming process causes both chemical and physical changes in the collagen. For example, during liming, deamination of the collagen amino acids such as glutamine and asparagine occurs with the evolution of ammonia, and the alkali labile cross-links between the chains are broken. For acid treatment, the gelatin has a higher isoelectric point (see Table 27.1) because the glutamine and asparagine residues remain in the amide form and are not converted to carbocyclic acid form as with the alkali treatment. As with liming, the

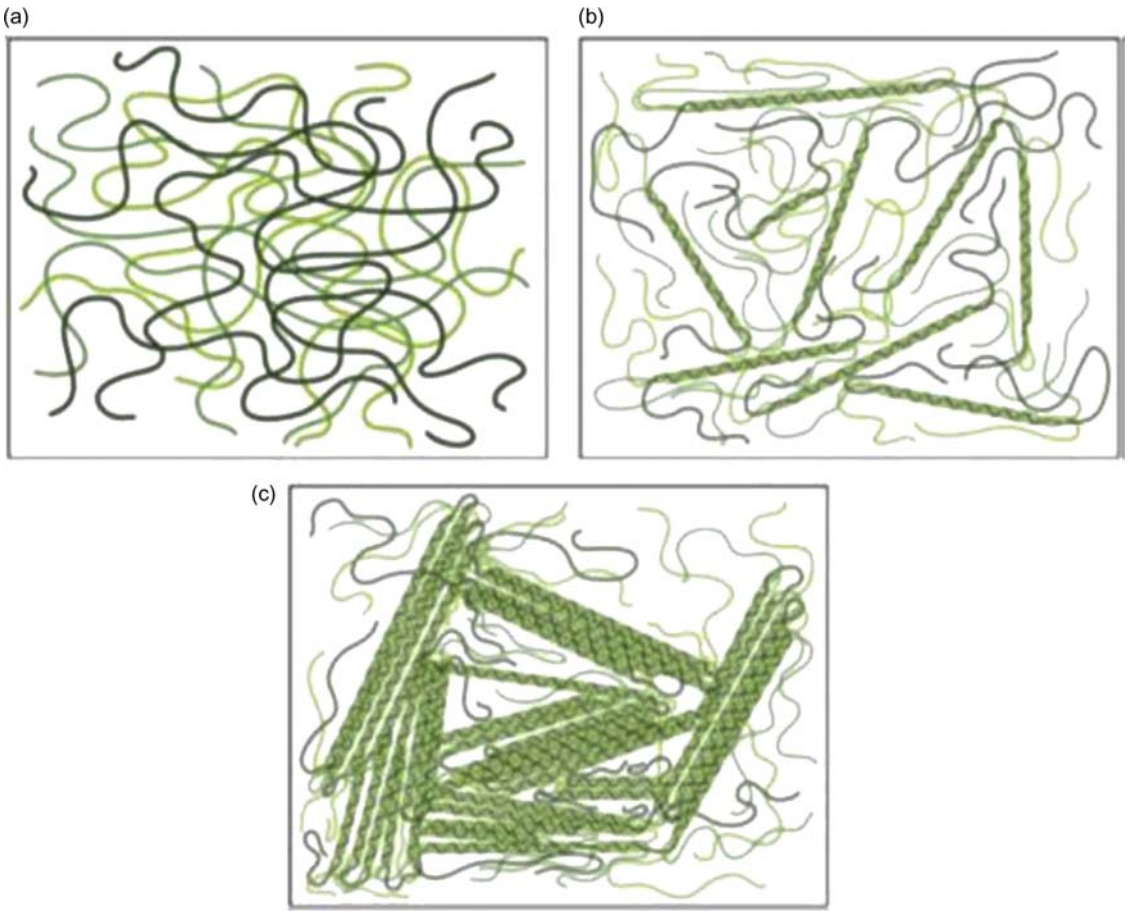


FIGURE 27.5 Changes in collagen structure that influence gelling properties. *Duconseille A, et al. Gelatin structure and composition linked to hard capsule dissolution: a review. Food Hydrocol 2015;43:360–76.*

acid treatment breaks up the collagen chains. The acid treatment is faster (on the order of 24 hours) than liming (8–12 weeks), but it has poor yields for many collagen sources such as bones and cattle hides.

After alkali or acid treatment, the collagen must be extracted from the liquor. The pH, time, temperature, and number of extractions (typically 3–6) vary with the product. The goal of extraction is to heat up the liquor under carefully controlled conditions to the point where the gelatin solubilizes; typically this is between 50°C and 60°C for the first extraction. The soluble materials are then filtered and or washed to remove the insoluble components. The soluble material is deionized, typically by ion exchange. The cleanup process is repeated several times until the material is very pure (Fig. 27.6). Then the soluble gelatin is concentrated and cooled until it gels and can be processed into a powder for shipping (Fig. 27.6).[10,11] The gelatin produced by the first extraction typically has a higher molecular weight, higher viscosity, higher gel strength, and lighter color, which is preferred for making capsules as compared to later extractions.

27.2.1 Capsule storage

A key component of gelatin is water. Newly made capsules will have a water content between 13% and 16%. If the amount changes, the capsule can become brittle and fracture when running on a commercial machine. Also, in our experience, if the water content changes, the dimensions can slightly change, leading to problems of capsule separation, rejoining, and solvent sealing. For example, when sealing fish oil hard-shell capsules, if the dimension change so the gap between the cap and the body increases, oil can leak into this space. This prevents the tight sealing of the cap to the body because the sealing solvent can't get into this space. Thus, capsule shells must be carefully stored. At a minimum, they should be stored in a tightly sealed box, and ideally, they should be stored in a temperature- and humidity-controlled room. For example, the Capsugel company recommends that for Coni-Snap capsules, the ideal storage conditions in the original corrugated box and tightly sealed polyethylene bag are 50% RH at 21°C. However, if the facilities are maintained

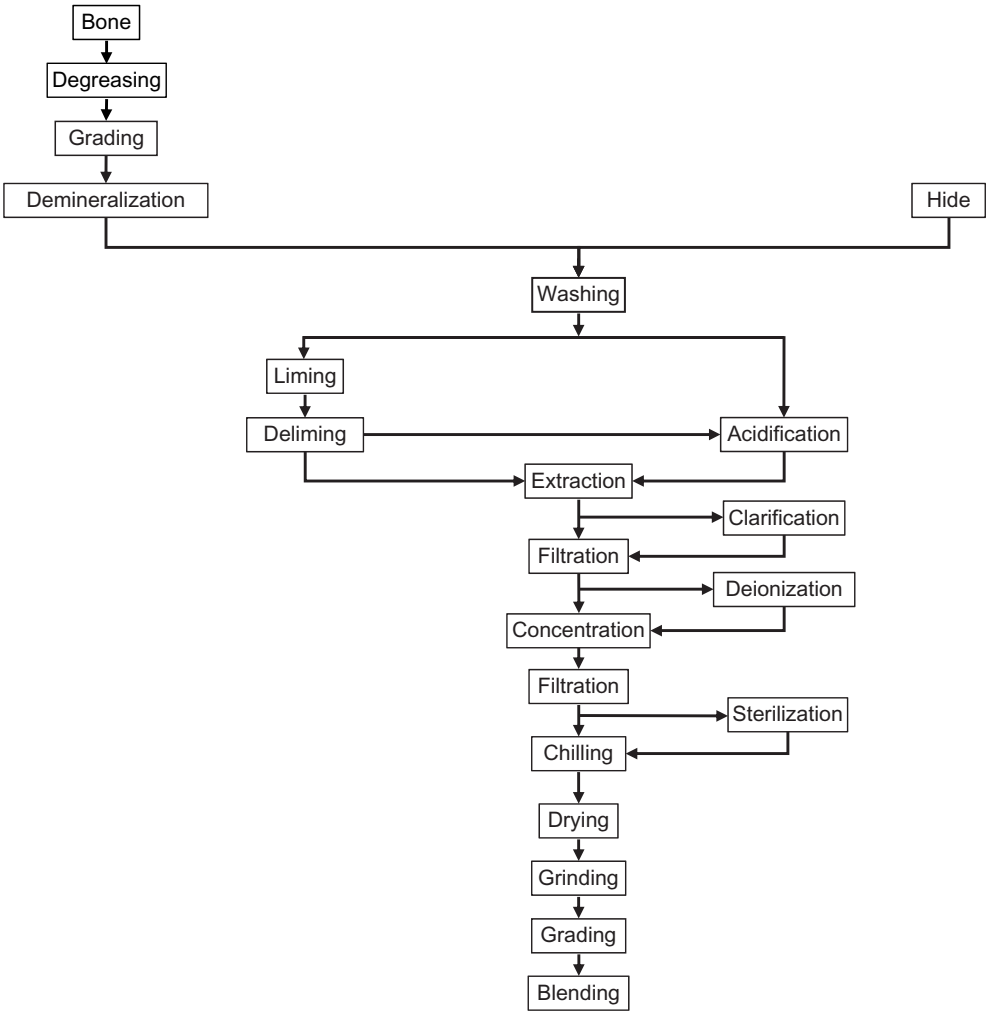


FIGURE 27.6 Gelatin manufacturing process. *Adapted from Jones.*

with a relative humidity in the range of 35–65% RH and temperatures in the range of 15–25°C, you should be able to keep them for the full 5-year shelf-life.

27.2.2 Gelatin cross-linking during storage

As discussed, gelatin undergoes a reversible thermal gelation due to the physical interactions such as hydrogen bonding between the protein chains (Fig. 27.4), and it is this gelation process that is in part responsible for the rapid drug release from a capsule. However, there are a whole host of irreversible chemical reactions that can occur between the gelatin protein chains that can cross-link the protein chains together.[12] When the extent of cross-linking becomes great enough, the gelatin becomes insoluble, and the capsule shell forms an insoluble film around the API. This swollen, rubbery, water-insoluble membrane is called a pellicle, and when it forms during dissolution testing, the drug release is greatly slowed, and the capsule fails to meet its dissolution specification. Common causes of gelatin cross-linking include aldehydes (in the active or excipients), peroxide impurities, storage in high heat and humidity, and rayon coilers.[12–15] When testing a moderately cross-linked gelatin capsule, it will fail dissolution due to pellicle formation; however, when tested in vivo, it may still have good bioequivalence compared to a noncross-linked capsule due to proteolytic enzymes in the gut.[12,16] These results indicate that the in vitro capsule dissolution test may not be predictive of what happens in vivo due to the digestive enzymes. Thus, the United States Pharmacopeia (USP) came up with a two-tier dissolution test, and if the capsules fail in the regular dissolution medium, they can be retested in medium containing proteolytic enzymes.

27.2.3 Capsule shell additives

Besides a structural polymer gelatin or hydroxypropyl methylcellulose (HPMC), there are other additives

in the capsule shell. Coloring agents are often added to gelatin because pure gelatin is clear, which allows light to contact the API and the patient to see the powder inside of the capsule. The most popular dyes added to capsule shells are synthetic water-soluble dyes (eg, azo, indigoid, quinophthalone, triarylmethane, and xanthene), pigments such as titanium dioxide and iron oxides, and natural dyes such as carotenoids.[17] When adding a coloring agent, one can either add a dye or a lake to the capsule shell.[18] In general, dyes impart a transparent or translucent color to the capsule shell, and the patient can still see though the shell, albeit a colored view. In contrast, pigments block the light and opaquely color the capsule shell, blocking light from reaching the inside of the capsule. Titanium dioxide (TiO_2) is the most commonly used opacifying agent. It has a white color and is mostly inert. TiO_2 can help the dye better opacify the shell; thus, they are often added together.

Printing inks are not really part of the gelatin, but they are materials applied to the capsule shell and could affect the product. Capsule printing is often done to meet Food and Drug Administration (FDA) product identification requirements, and the printing ink colorants must meet the same regulatory requirements as those for the capsule shell.[1] Generally, inks are made up of insoluble colorants dispersed in a volatile solvent and a film former.[19] The colorants must be insoluble, or when the solvent evaporates, it will bleed, making the printing look blurry. The coloring agents can either be added as a pigment like the iron oxides, which are insoluble but limited in colors, or as lakes. Lakes are made by precipitating a dye onto an aluminum hydroxide or titanium dioxide carrier, thus, making them insoluble.[18] Lakes have the advantage of coming in a wide range of colors and are insoluble. A common ink film former is shellac, which has broad acceptance in the major pharmacopeias and adheres well to gelatin. Typical solvents used in the capsule printing inks include ethyl alcohol, isopropyl alcohol, *n*-butyl alcohol, and water. Obviously, one has to be sure that the residual solvent levels don't exceed regulatory limits. Fortunately, the amount of ink applied is small, so this is typically not a problem if dried properly.

A warm gelatin solution can support rapid microbial growth, so care must be taken to keep the total microcount low when working with these solutions. A properly stored capsule will have a water content in the 13–16% range, which is low enough to not support microbial growth. Thus, preservatives are not used in gelatin capsules manufactured in Europe and in the United States. However, in older products, preservatives such as methyl and propyl parabens and sodium metabisulfite were added to gelatin solutions to control microbial levels during production.[7]

Sometimes processing aids can be in a capsule shell. For example, sometimes a surfactant such as sodium lauryl sulfate can be used in capsule shell manufacturing. Also, there can be proprietary pin-release agents used to help the capsule shell be stripped from the pin after dipping. Also, things such as silicon dioxide have been used in gelatin manufacturing. Some additives are allowed in the USP, while others are disclosed with a confidentially agreement with the capsule shell manufacturer. If there are stability problems, gelatin additives should be considered when trying to solve the problem.

27.2.4 Mad cow disease

In the 1980s and 1990s, BSE (also known as mad cow disease) was found in cattle in Europe, and in the United Kingdom people were actually infected, most likely from eating animal products containing brain or spinal cord parts. While the process of producing gelatin is very harsh with either an acid or alkaline treatment and reduced the infectivity to undetectable levels, there is still a very slight possibility that the prion that causes BSE could survive the process.[6] Given that prions are difficult to detect, suppliers of gelatin must have preventative measures in place so their products do not contain prions. These measures include obtaining gelatin from certified BSE-free production areas. Only healthy slaughtered animals can be used, and the gelatin must not contain animal parts exposed to neurological tissues such as the skull and backbone vertebrae. The regulations for certifying that gelatin is BSE-free varies from country to country. For example, Japan and Argentina require only hide gelatin. Thus, make sure that suppliers can supply a BSE-free certificate, as this will be needed for release. Typically manufacturers have these on their websites for download, and the certificate should be kept as part of the batch records for clinical studies and product release to the market.

27.3 CAPSULE SHELL MANUFACTURING

The first step in making a hard-shell capsule is to produce a well-mixed dipping solution; the history of capsule making is discussed by Jones et al.[7] In addition to gelling polymers, gelatin, or HPMC, the solution could contain other additives such as coloring agents.[7] The key quality attributes of the gel solution include viscosity and gelation temperature. To achieve the necessary attributes, the gelatin solution is made from a combination of Type A and B gelatin for the best combination of shell strength and clarity.

Capsule shells are made by dipping pairs (body and cap) of room-temperature stainless-steel pins into a heated gelatin solution. Because the pins are below the gelling temperature, the gelatin begins to form a thin gelatin layer or film on the pins (Fig. 27.7).[20] After a short time in the gel solution, the pins are removed and rotated several times to evenly distribute the gel over the pin. Once the gel is evenly distributed on the pin, a blast of cool air is used to set the gelatin on the pin. At this point, the gelatin is dried, and the pins are sent through a long series of continuous ovens until the moisture content is at the desired level (Fig. 27.8). The drying is primarily done using dry, slightly warmed air to avoid softening or melting the gelatin in the drying oven; recall the gelling temperature of gelatin is about 30°C. After the gelatin is dried, the capsule is striped off the pin and trimmed to the proper length. Once trimmed, the two halves are put back together in the prefit position. At this point, printing is done if needed before packing in cartons for shipping.

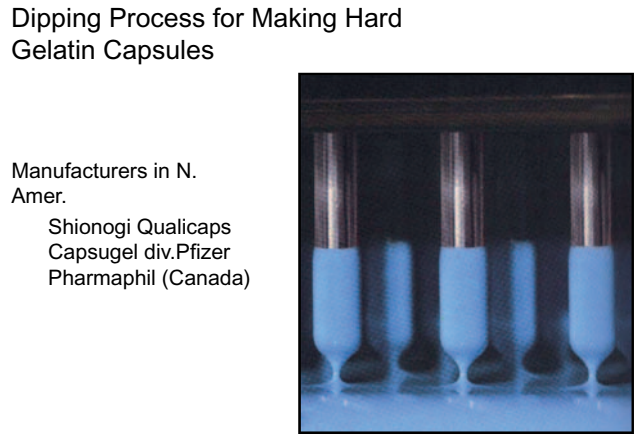


FIGURE 27.7 Dipping pins in a gel solution.

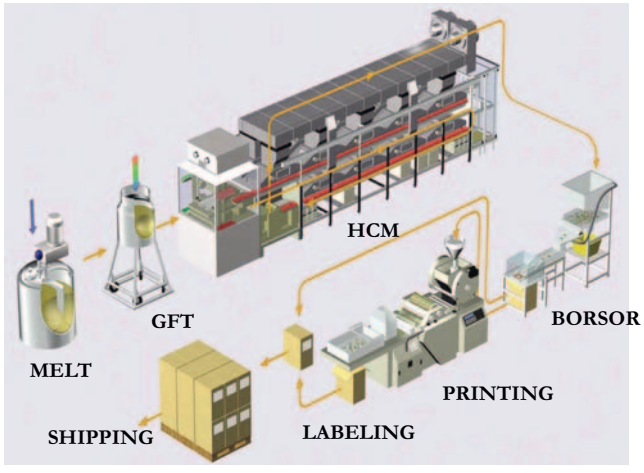


FIGURE 27.8 The hard gelatin capsule manufacturing process. *Source: Dennis Murachanian, Capsulgel.*

27.4 ALTERNATIVES TO GELATIN

Issues with BSE, religious dietary restrictions with pork-based products, consumer preferences for healthy veggie products, and animal rights concerns have led to the demand for nonanimal-based capsule shells. Given the exacting processing requirements where capsule shells have to run smoothly on the same commercial filling machines as the gelatin capsules and the drug delivery requirements of capsule shells, it has been hard to find cost-effective alternatives to gelatin capsule shells. Despite these difficulties, there are now a range of gelatin alternatives readily available on the market, but these products were not easy to develop.

The most common alternative to gelatin is hypromellose in USP official nomenclature, also known as HPMC. All major suppliers of capsules now produce HPMC capsules. HPMC is a cellulose ether of vegetable origin, which eliminates concerns with religious restrictions and dietary preferences against animal products. Like gelatin, HPMC has a long-use history in the pharmaceutical industry for things such as a wet granulation binder, layering binder, controlled-release matrix former, film coating material, and suspending agent. It has an excellent safety profile, is considered GRAS (generally regarded as safe), and is listed in all the major pharmacopoeias.[5] HPMC has lower water content, on the order of 4–7% compared to gelatin, which is in the range of 13–16%. This reduces problems with brittleness due to loss of water because water does not plasticize HPMC as much as it does gelatin, and a lower water activity level generally is good for moisture-sensitive drug stability.[1] There are also differences in water and oxygen permeability, which is important for drug stability.[1,21]

HPMC capsules are made by the same dipping process as gelatin. To form an HPMC capsule, the HPMC solution must rapidly form a film on the dipping pin. Currently, this film can be formed either by thermal gelation or gelling via a gelling system. Unlike gelatin, HPMC undergoes reverse thermal gelation; in other words, HPMC forms a gel at higher temperatures, whereas gelatin forms a gel at lower temperatures.[22] For cellulose-based polymers such as HPMC, it is not unusual for them to be more soluble at lower temperatures, because at higher temperatures, the hydrophobic interactions predominate more than the hydrophilic interactions, and the hydrophobic interactions lead to aggregation between the polymer chains. As the number of aggregates goes up, the gelation occurs. This transition temperature can also be affected by environmental factors such as ionic strength.[23] Silva has shown

that the sol-gel transition temperature occurs at about 55°C; however, this transition temperature is affected by environmental factors and differences in the HPMC polymers such as the molecular weight and the degree of substitution.[22,23] Thus, to make HPMC capsules on the same dipping pin system, the HPMC solution is kept cool at room temperature, and the pins are heated to above the gelation temperature.[5] The hot pins are dipped into the cool HPMC solution, removed from the solution, spun on the pins to form a uniform film, and then dried, much like the gelatin capsule-making process.

The disadvantage of thermal gelation using a cold HPMC solution and hot pins to make the capsules is that very expensive capsule-making equipment must be modified or developed to accommodate these differences in processing. Another approach is to add gelling agents to the HPMC solution, and these gelling agents have the property of being a solution at cold temperatures and a gel at higher temperatures. This allows the use of the same cold solution/warm pin equipment when making HPMC capsules.[5] Currently there are two commercial polymer gelling agents: kappa-carrageenan and gellan gum. Kappa-carrageenan is a structural polymer obtained from red seaweed. It is a sulfonated polysaccharide that is widely used in the food industry as a gelling and thickening agent, and it has an excellent safety profile. Kappa-carrageenan has the unique feature that it forms helical structures at room temperature, and the gelation occurs via interactions between the helixes. Being a sulfonated polymer, the sulfate group is unionized at the typical working pHs of κ-carrageenan, which makes the polymer very sensitive to ions, and small changes in ionic strength can affect the gelling behavior of κ-carrageenan.[24,25] This sensitivity is used to control the gelling temperature of κ-carrageenan, so it can be processed on commercially available dipping machines.

The other polymer used commercially as a gelling agent for capsule manufacturing is gellan gum. Relative to κ-carrageenan, gellan gum is new polymer that was only discovered in the late 1970s. It is produced by the bacterium *Sphingomonas elodea*. Gellan gum is a linear anionic polysaccharide that has one carboxyl side group and one 0-acetyl substituent per chemical repeat unit.[26] The carboxylic group, which can ionize at the typical working pH of the polymer, makes the gelation temperature sensitive to environmental factors such as ionic strength. Like κ-carrageenan, the ionic strength and the pH of the gelling solution can be controlled so that gellan gum gels at a good temperature for making capsules on commercially available dipping machines.

Examples of commercially available HPMC capsules made by thermal gelation include Vcaps Plus, also called tg-HPMC. Commercial capsules made with κ-carrageenan include QualiV, also called c-HPMC, and commercial capsules made with gellan gum include Vcaps, also called g-HPMC.[5] A critical issue with capsule performance is how these three types of capsules compare to each other and with gelatin. The key properties such as water uptake, mechanical strength, color-hiding power, stability, esophageal transit time, in vitro dissolution, and in vivo bioequivalence of these different types of capsules have been extensively studied.[5,27–29]

These studies and others are hard to summarize because the results depend upon the type of drug and the test conditions used, but in general, they found that the performance was similar, but there were important differences between the different types of capsules shells. Thus, without some type of evaluation, one cannot assume that the capsule shells types are interchangeable because there could be conditions where the differences could affect in vivo performance and patient outcomes. A purchasing agent should not make the decision to switch from one type of capsule shell to another, as this should be done within the confines of a company's change control procedures.

Currently, HPMC capsule shells have been developed to the point of being commercially available from all the major suppliers; however, in the future, there will be many new developments in capsules shells. Some of these new technologies include pollen capsule shells, starch hydrolysate capsule shells, HPMC softgel shells, and capsule shells that are designed for modified release; for example, there are enteric capsule shells and controlled-release capsule shells for colonic delivery. Because these new innovations are not as widely available as HPMC, they will not be discussed further, but they have been reviewed by Stegemann.[5]

27.5 HARD SHELL

The most commonly available capsule shells are the hard gelatin capsule shell. This shell comes in two parts: the body and the cap (Fig. 27.9). These capsule shells are made of a hard gelatin or HPMC, which is ridged and can be handled and run though processing equipment. This is in contrast to soft gelatin shells, which have to be formed and are typically used for liquids and semisolids. Originally, the cap and the body just fit together without a locking mechanism, but this lead to problems, as sometimes the cap and the body separate during handling and shipping; thus, most capsule shells today come with a locking mechanism that keeps the cap and the body from separating. The Coni-Snap capsule shell is one such example, as you can see in Fig. 27.10. There is a locking ring that, once engaged, holds the cap and the body together.

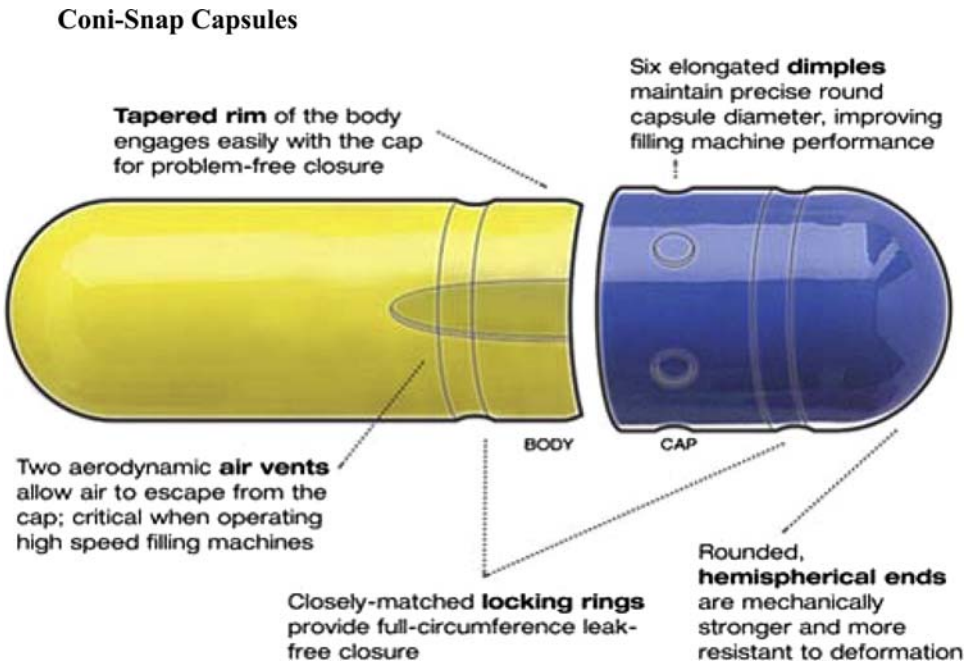


FIGURE 27.9 Capsule shells showing features.

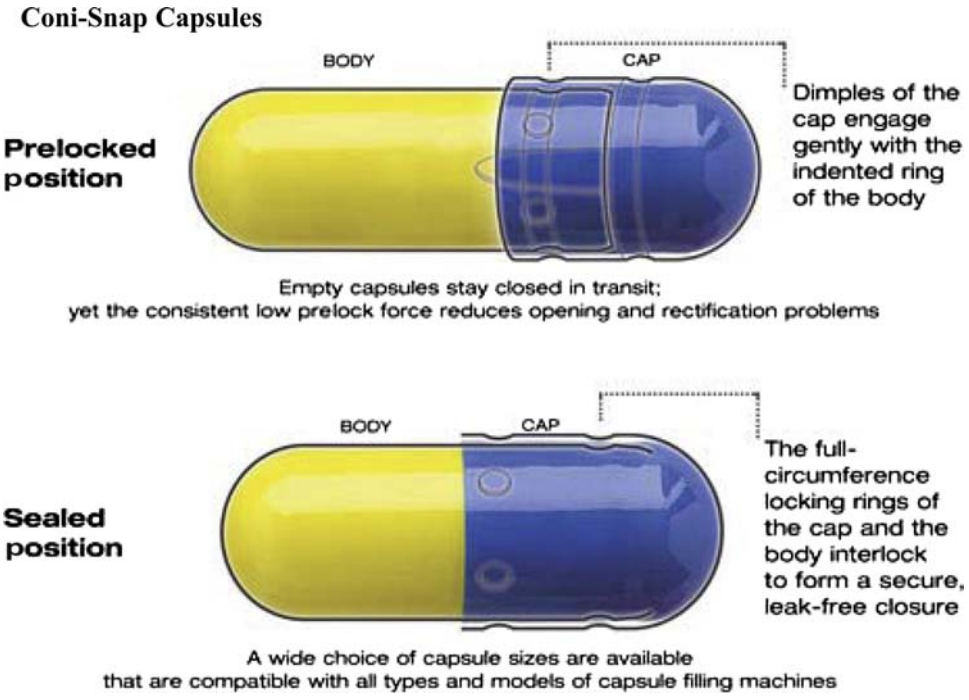


FIGURE 27.10 Capsule shells.

As shown in Fig. 27.10, the shells come from the manufacture in the prefit position, and after filling, they are pushed together to lock. Note that in this context locked is not the same as sealed.

Modern capsule-filling equipment can fill up to 200,000 capsules per hour. For a capsule shell to run on such a high-speed machine, it places a lot of performance requirements on the capsule shell. Things such as the mechanical strength of the gelatin film, dimensional accuracy of the shell, the ability to vent air during high-speed closure, closure-locking mechanism, prefit locking strength, the ability to align for rectification, and

closure are all essential qualities that are needed to run on a high-speed machine. How these features are engineered into a capsule shell is part of a company's patent portfolio and used as special features to sell capsule shells.[1] Traditionally, three of the biggest manufacturers of capsules were Capsugel, who made the Snap-Fit and Coni-Snap capsule shells; Pharmaphil (now part of Qualicaps), who made the Lox-it capsule shells; and Shionogi Qualicaps (now called Qualicaps), who made the Posilok capsules. This is not a complete list of all manufacturers current or past. During the past decade, a lot of companies in China, India, and other parts of the world have started to manufacture capsule shells; they are too numerous and transitory to review here.

27.5.1 Capsule sizes

Capsules come in standard sizes (Fig. 27.11). These sizes range from 000 to 5, which have volumes from 1.36 to 0.13 mL, respectively. To estimate the fill weight of these capsules, one can use the basic bulk or tapped density if you compress during filling. The bulk density will give you an approximate fill weight, which can then be adjusted by trial and error. A more quantitative method for determining the fill weight is to use the method developed by Khawam.[30,31]

27.5.2 Use in preclinical and clinical studies

For clinical research, when one wants a blinded placebo or to make a comparison to a particular formulation or existing product, it can be very difficult because making an exact look-alike copy of a product can be very difficult and time consuming. Thus, capsule manufacturers have developed special-sized capsules with a wide body specifically designed to make it easier to be directly filled with an intact tablet (Fig. 27.12).

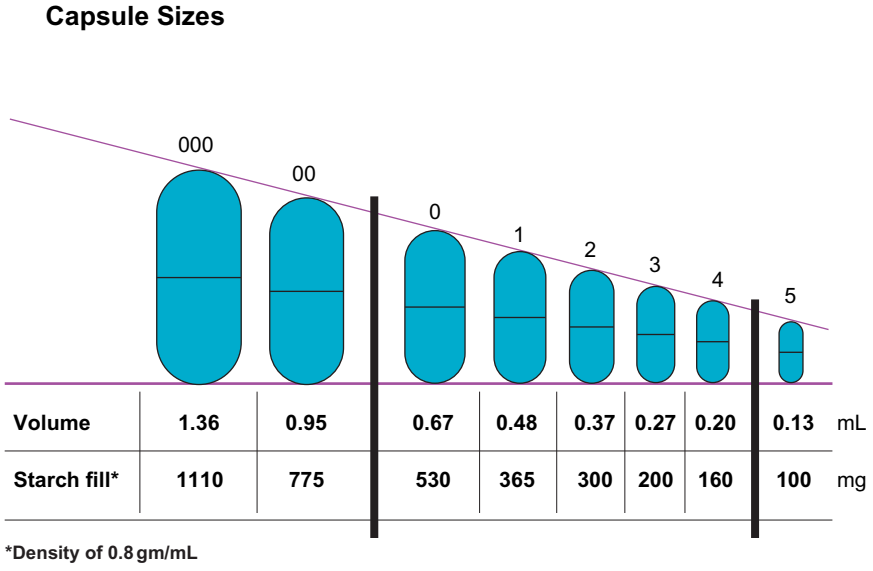


FIGURE 27.11 Standard capsule sizes. *Source: Dennis Murachanian, Capsulgel.*

FIGURE 27.12 Double-blinded capsules for clinical studies.

There are eight different sizes of double-blind (DB) capsules: AAA, AAel, AA, A, B, C, D, and E; see the Capsugel website for size details. These capsules can be filled by a compounding pharmacist. This can avoid the expense of good manufacturing practice (GMP); however, many compounding pharmacists lack the capacity to support even small clinical studies. The double-blinded capsules have some unique features that make it hard to break the blind. For example, the body is extended, and the cap fits into the body, making it very difficult to pull them apart without destroying the capsule, which makes it almost impossible for the patient and curious nurses or physicians to break the blind without obvious signs of damage, which can alert researchers to this possibility. Obviously this is not a foolproof method of blinding, but it will at least stop the casual patient or health care worker from breaking the blind.[32]

27.5.3 Animal testing

Administration of an API to an animal is one of the important initial steps in preclinical research. If you are administering the drug to an animal such as a dog, pig or primate, you can use human dosage forms, but if you need to administer an API to a rodent, then human dosage forms are too large. For preclinical research, the API is typically administered as a solution, emulsion, or suspension via gavage; however, some APIs may have poor stability, and a liquid preparation would not have the needed shelf life for testing. Some APIs may be very difficult to formulate as a liquid predations, and if the study questions have to do with the API in the solid state, then a liquid preparation will not answer your questions. For these situations, capsules have been specifically developed and designed to be delivered to mice, hamsters, rats, and guinea pigs (Fig. 27.13). In addition, they have developed specially filling apparatus and tools for capsule administration, also shown in Fig. 27.13. These capsules are hand-filled one a time (see Section 27.6.1 for a discussion of hand-filling), so their use is limited to preclinical studies. Also, these capsules can be filled without excipients, which is an advantage when compactability and stability are a concern. On the other end of the spectrum are large specialty capsules for livestock. For example, there are capsules that can deliver gram quantiles of medication orally, rectally, and vaginally to cattle.

27.6 CAPSULE FILLING

As mentioned previously, capsules are shipped with the cap and the body prefit together. Thus, to fill a capsule, the basic steps are: separate the cap from the body, fill the formulation into the body, put the cap back onto the body, and lock the cap and the body together. One of the beauties of capsule filling is that it is very flexible in terms of scale. Capsules can be hand-filled one at a time, as done in a compounding pharmacy, or they can be hand-filled using a jig, a semiautomated capsule-filling machine, or filled on commercial-scale machines making over 150,000 capsules per hour.

The primary differences between small-scale filling and commercial production is in how the powder is transferred into the capsule body. Small-scale production is done by directly filling the powder into the capsule shell and relying on the bulk/tapped density of the powder to get the correct dose for the volume of the capsule shell used. Capsules can also be made by direct weighing of the API into the capsule body. For commercial production, the capsule is filled in a process called *indirect filling*, which is done by forming a plug that is transferred into the capsule shell. Forming a plug has several advantages over direct filling. The plug is made by compressing the powder into a cohesive mass, which holds together during the transfer to the capsule body. This can be done by multiple

FIGURE 27.13 Capsule used for administration to rodents.

tamping on a dosing-disk machine or by compression in a dosator. The advantage of forming a plug is that the powder can be compressed to higher density, which allows for higher capsule fill weights, which given the patient preference for swallowing smaller capsules can be important. This is very important for botanical products that typically have very low bulk densities and if not compressed can lead to the patient having to swallow very large capsules. In addition, the machine operating parameters of the dosing-disk or dosator machine can be adjusted to accommodate the natural variation of the formulation components to product more consistent fill weight, thus, improved dose uniformity.

27.6.1 Hand-filling capsules

Hand-filling is the easiest and the slowest method. It has the least formulation requirements and can accommodate just about any API that is compatible with the capsule shell materials, but this method is not scalable to faster modes of production. Hand-filling is typically done to prepare a small number of capsules for toxicity and bioavailability studies in animals and humans. Using a capsule-filling jig, one can produce enough capsules to support a phase I study or clinical research studies. Hand-filling can either be done one capsule at a time or by using a capsule-filling jig.

Hand-filling one capsule at time is what is commonly done in a compounding pharmacy, where you form a little pile of powder on a pill tile and then push the inverted capsule body into the powder blend until the capsule body is filled up and has the desired weight. Alternatively, you can weigh the API directly into the capsule body, but this is very slow process. In recent years, an automated weight system has been developed for direct filling of the API into a capsule shell; the system is called the Xcelodose system, which is described in Section 27.6.3. The capacity of hand-filling is about 5–20 capsules per hour depending upon the material being filled and operator skill.

To increase manual production rates from that of a pill tile, one can use a capsule-filing jig, also called a hand-filling capsule machine (Fig. 27.14), such as those sold by CroPharm Inc., Torpac Inc., and Capsugel Inc. Typically, these jigs can make from 100 to 500 capsules per fill (Fig. 27.14). The basic steps are to (1) rectify the capsules, (2) separate the cap from the body, (3) fill the powder, (4) join the cap and the body together, and (5) eject the filled capsules from the jig. These steps are shown in Fig. 27.14. Because the cap and the body have slightly different dimensions, the basic idea of the hand-filling jig is to use a series of plates with different sized holes to independently manipulate the cap and the body for all the filling steps. To rectify the capsules, they are randomly poured onto a tray that has a series of slots, one slot for each capsule. The middle of the slot is bigger than the ends, so the capsules fall into a hole body first. At the bottom of the hole are four plates with a hole for each capsule. The top plate has a hole that allows the body to pass through but not the cap. Below the top plate are three plates (Fig. 27.14). The middle plate is on an eccentric cam, which moves the middle plate relative to the other two plates in order to lock the capsule body in the three plates. Once the capsule bodies are locked in place, the top plate is removed along with the caps, resulting in the separation of the cap from the body. The capsule bodies can be pushed flush with the top of the three lower plates so when a powder mass is poured onto the jig plate, it can be evenly distributed to each capsule. Often during the filling of the capsules a tamper is used to slightly compress the powder. This increases the bulk density of the fill and helps distribute the powder into each capsule. Once the capsules are full, they are rejoined by putting the top plate back on and squeezing the stand so the capsule bodies are pushed up through the three middle plates and into the top plate where the caps are held. Then the top plate is removed, the capsules are manually pushed into the locked position, and they are dumped into a tray for packaging or further processing.

27.6.2 Liquid filling two-piece capsules

For some drugs, a liquid fill can be better than a dry powder; for example, for drugs with low solubility, the delivery can be enhanced with a self-emulsifying drug delivery system.[33] Also, with highly potent, for example, hormones, or cytotoxic drugs where worker exposure to dust is a problem, a liquid fill can provide better containment, and with low-dose drugs, content uniformity can sometimes be easier to maintain in a liquid formulation rather than a hard-to-mix powder.[34] Typically, liquids are filled in soft-shell capsules, but with special modification of traditional filling equipment, liquids and semisolids can be filled in two-piece capsules. Liquids can also be hand-filled into two-piece capsules for feasibility studies and clinical studies.

For liquid fills, all the steps are the same as with the powder fills, but instead of filling a powder into the capsules, a pipet or liquid manifold is used to fill each capsule with a liquid (Fig. 27.15). For liquids, the cap and the body must be sealed to prevent leaking. This is very important for oils such as omega-3 fatty acids that turn brown when oxidized, which is a very obvious defect on the surface of a capsule shell. Thus, once the capsules have been filled, it is important to quickly

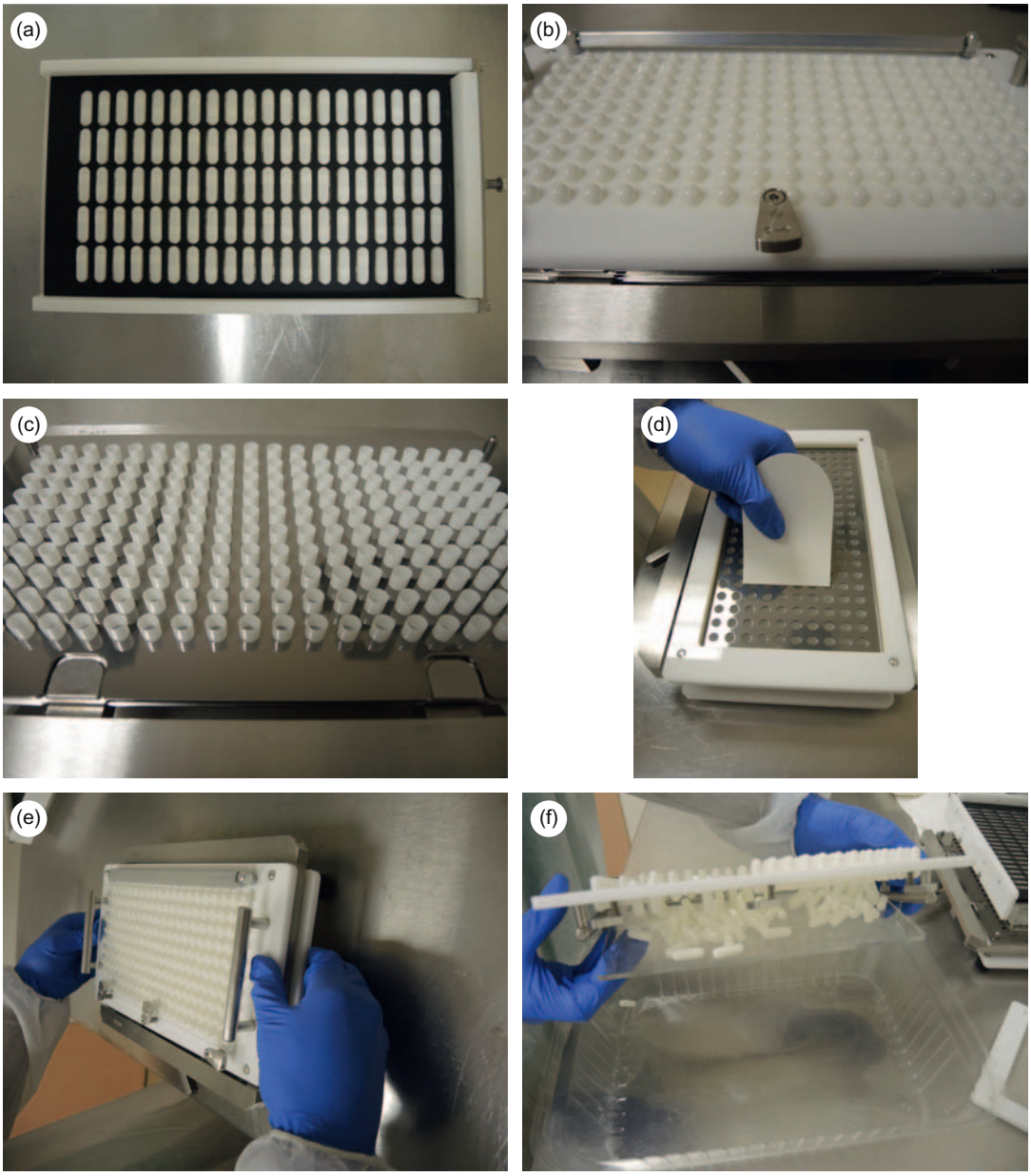


FIGURE 27.14 Steps for using a hand-filling jig: (a) rectification of capsules, (b) capsule separation, (c) separated capsules ready to fill, (d) using plastic scraper to distribute powder in capsule shells, (e) closing capsule shells, and (f) emptying filled capsule shells.

seal them. Capsules can be sealed either by banding with heated gelatin or a hydroalcoholic solution fusion process.[35] In addition to liquid filling for over-the-counter product, capsules should be banded so as to be tamper evident.[36]

To seal a two-piece capsule using a hydroalcoholic solution, the solution is either sprayed onto the capsule shells in a pan coater or, if hand-sealing, via a swab. When the liquid comes in contact with the gap between the cap and the body, the liquid surface tension wicks the solution into the gap, and then gentle heat is applied, typically between 40°C and 50°C, for a minute. Typically a 50:50 water-ethanol solution is used. The water in the alcohol solution softens the partially dissolving the gelatin, and when heated, the two pieces fuse together. Then the water evaporates, and the gelatin returns to its dry state, but the two halves are bonded together. If pure water is used, too much softening occurs. The capsule may develop ripple defects, and sometimes the gelatin bubbles when heated. If too much water is applied, this can lead to twinning (when two capsule shells stick together), and at the point where the shells are conjoined, there is typically a spot defect in the gelatin. To band a capsule requires special

FIGURE 27.15 Liquid filling of hard gelatin capsules.

equipment that heats up the gelatin and then applies a thin band around the joint between the body and the cap.[34,35] When liquid filling two-piece capsules, it is best to use capsule shells specifically designed for liquid filling. These capsules have features such as no vents and are designed for improved sealing of the liquid or semisolid. Liquid filling can also be done on a commercial scale, but this is beyond what will be covered here, and the interested see reader can see Cole et al.[34] and Niederquell.[37]

27.6.3 Powder in capsule automated filling

It is well known that only a small fraction of the compounds synthesized make it to the market. Given the extreme risk of developing new compounds, the goal of most companies is to make a go-no-go decision as soon as possible for any compound in development. A pivotal point in the development process is the first studies in humans, and to conduct these studies, one needs a formulation to administer to the patient. One way to develop a formulation rapidly is to directly weigh only the API into the capsule, which is called the powder-in-capsule (PIC) method. With only the powder in a capsule, you don't have to worry about adding an excipient to make the capsules. For example, you don't have to worry about lubricants and flow aids in the formulation, and consequently, you don't worry about excipient compatibility and how much to add to make the process run well.

To meet this need for rapid PIC formulation manufacture, companies such as Capsugel Inc. have developed an automating weighting system that can fill between 100 μg to 100 mg per capsule, and according to the company, it can do this with a 2% RSD. The trade name for this system is Xcelodose. The system accommodates up to a size 00 capsule, and the automated version of Xcelodose can fill up to 600 capsules per h. As discussed, hand-filling via direct weighing is very slow, and without an automated system, hand-weighing would not be practical for a phase I study. A more detailed description of the Xcelodose system and its operation for filling capsules to put in a dry powder inhaler can found in the review by Edwards.[2]

27.6.4 Semiautomatic

The next level of filling is a semiautomatic capsule-filling machine, shown in Fig. 27.16. Conceptually, this is the same as the hand-filling jig, that is, with no plug formation. The only differences are that a few more steps are automated and scaled up, and the powder is fed into the capsule shells using an auger rather than by manual feeding (Fig. 27.16). Some semiautomatic-style systems use vibration to fill the powder into the capsules, but to the best of the author's knowledge, these systems are not that popular and won't be discussed further. In the old days, semiautomatic machines were commonly used for commercial production, but over the years, the plug-forming indirect-filling machines have largely replaced the semiautomatic machines for commercial production, and today these machines are typically used for small-scale, flexible manufacturing of batches for feasibility studies and early-phase clinical trials.

With a semiautomatic capsule-filling machine, the process of capsule rectification is automated (Fig. 27.16). With these systems, the capsules are randomly fed into a slot, and then a finger pushes in the middle of the prefit capsule. Because the body is smaller, it moves first, which orientates the capsules so they are all facing in the same direction (Fig. 27.17). Once the capsules are all aligned, they are pushed into the dosing rings. As with the capsule-filling jig, when the rings are manually separated, the cap is separated from the body. After separation, the body ring (Fig. 27.16) is placed under an auger-driven hopper and rotated on a turntable. To fill the capsules, the body ring is rotated, and the hopper auger is turned on so the powder is fed from the hopper at a fixed rate as the body ring rotates. The fill weight is determined by a combination of the turntable speed and the auger speed. The faster the turntable rotates, the lower the fill rate. As there is less time under the auger, which dispenses powder at a fixed rate, the faster the auger rotates, the higher the fill weight, as a faster auger speed pushes the powder out faster into the capsule bodies. Once all the capsule have been filled, the rings are pushed back together to rejoin the cap and the body, and finally the capsules are ejected from the rings (see Fig. 27.16).

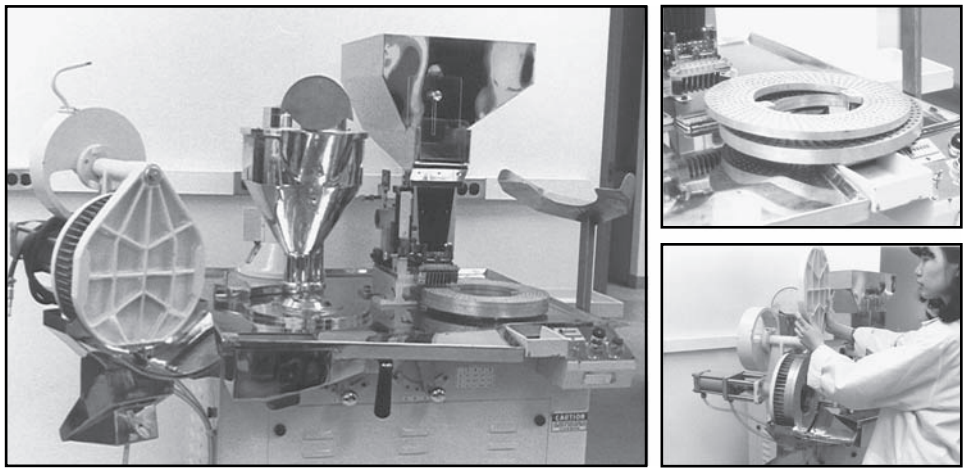


FIGURE 27.16 Semiautomated capsule-filling machine (eg, Colton Type 8).

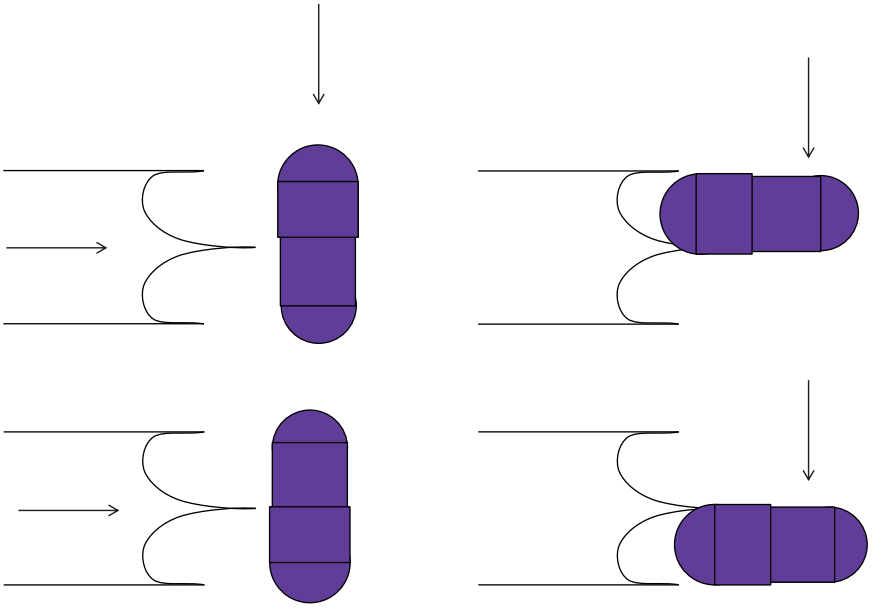


FIGURE 27.17 Capsule rectification.

Typically these machines can produce up to 10,000–15,000 capsules per hour, depending upon the ease of formulation handling and operator skill. The formulation requirements are essentially the same as hand-filling, but the powder has to flow out of the hopper with the aid of an auger, so the formulation can't segregate, and it needs to be relatively free flowing. Lubricants and flow aids may be necessary if discharge from the hopper is problematic. If the material is known to be sticky, that is, adhere to metal, then formulation changes must be made for the formulation to run well. Formulation strategies for API sticking to capsule-filling machine parts are similar to those for dealing with sticking and picking in tablets.[38]

27.6.5 Commercial production methods

The indirect method of capsule filling involves forming a plug and then ejecting the plug into the capsule body. This method of capsule filling is used in the vast majority of commercial filling operations. There are two principal methods for making a plug, the dosing disks and the dosator methods, and these commercially available capsule-filling machines can make 150,000 or more capsules per h. To reliability fill at high speeds requires a very good formulation. It is a misnomer to say there are no formulation requirements for a capsule because all you have to do is get the drug into the capsule shell. While this may be true for hand-filling or

slow filling on a semiautomatic capsule-filling machine, for high-speed production, this is not the case. In the next sections, we will discuss in general terms the steps of capsule filling for the dosing-disk and dosator method of filling followed by a discussion of capsule formulation. For a history of capsule filling and a detailed discussion of capsule-filling machines, see Lightfoot.[39]

27.6.5.1 Dosing disk

The dosing-disk method of filling uses a disk with holes and tamping pins to form a plug (Fig. 27.18). This type of machine is exemplified by the Bosch (formerly Hofliger Karg) GKF models and the Harro Hofliger (H&H) KFM models. The dosing-disk setup, shown in Fig. 27.18, has the powder bed above the dosing disk, and it uses tamping pins to push the powder into holes in the dosing disk. A typical dosing disk has from 12 to 100 holes drilled in the dosing disk. To simplify the discussion and match the discussion to the illustrations, I will refer to the H&H machine, which has a dosing disk with 12 holes. This discussion can be easily extended to other machines such as the Bosch GKF machine, which has 72 holes, so the same things are happening, but instead of 12 holes, there are 72 holes in six groups, and the dosing disk rotates 60° instead of 90° like the H&H machine. The matching tamping pin assembly has a tamping pin to match each hole. These are mounted onto a single assembly, with all tamping pins going up and down simultaneously. Typically, the tamping pin assembly tamps, that is, goes up and down, 2–5 times for each hole to form a plug. Then the disk rotates to the next set of tamping pins and compresses the plug three to five times more. With each tamp, the capsule plug become taller and denser. After three sets of tamps, the dosing disk rotates to the final hole, which is open on the bottom so that the capsule plug is unsupported. The capsule body is below the final position of the dosing disk, and when tamping pin comes down, it ejects the plug into the capsule body. The tamping pins are spring loaded to prevent overloading the capsules, which are not designed to take large forces. Typical tamping forces range from 20 to 200 N (see Fig. 27.19). For a more detailed description of this process and the monitoring of plug formation, see Shah and Augsburger.[40] To have a consistent fill weight, it is essential to keep the powder bed at a constant bulk density. Thus, there are typically moving baffles in the powder bed that sweep across the dosing disk to keep the bed a uniform density. The primary factors controlling the fill weight are the thickness of the dosing disk, the thickness of the powder bed, and the tamping pressure.

27.6.5.2 Dosator

A dosator, shown in Fig. 27.20, is made up of a dosing tube and a piston that slides up and down in the dosing tube. Typical dosator machines include Zanasi, MG2, Dott. Bonapace, and Macolar machines. To make a plug, the dosing tube is inserted into a powder bed, and the piston compresses the powder to form a plug. Then the dosing tube is withdrawn from the powder

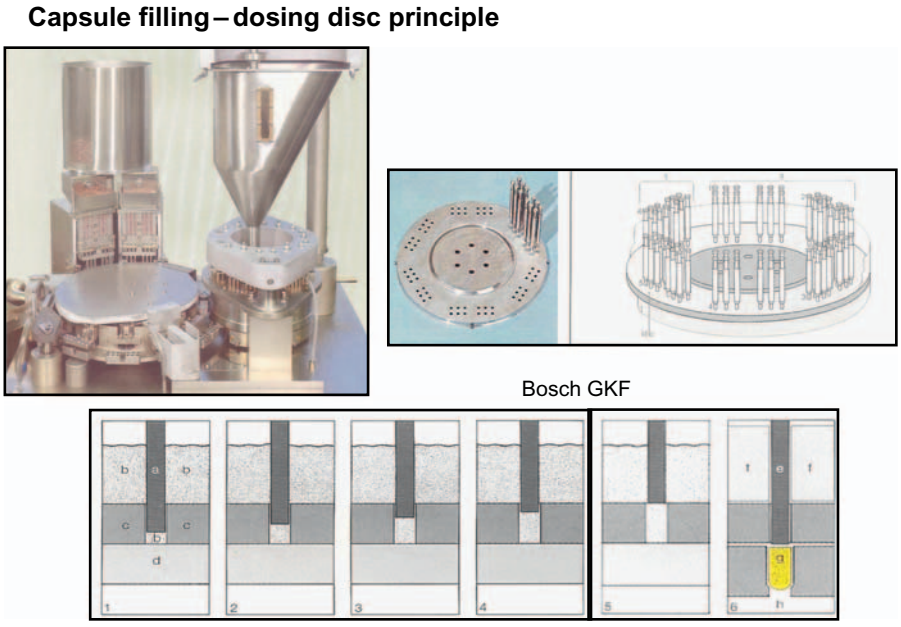


FIGURE 27.18 Schematic of Bosch dosing-disk capsule-filling machine.

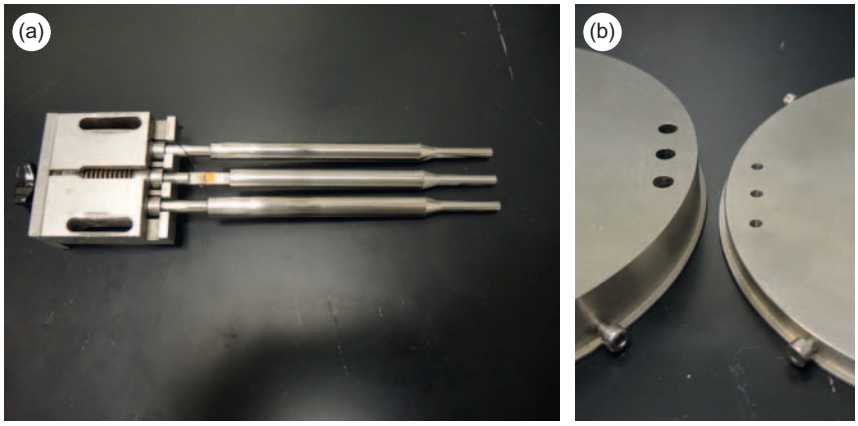


FIGURE 27.19 (a) Spring-loaded tamping pin with strain gauge instrumentation and two different dosing disks showing size differences. (b) Spring-loaded tamping pin with strain gauge instrumentation and two different dosing disks showing size differences.

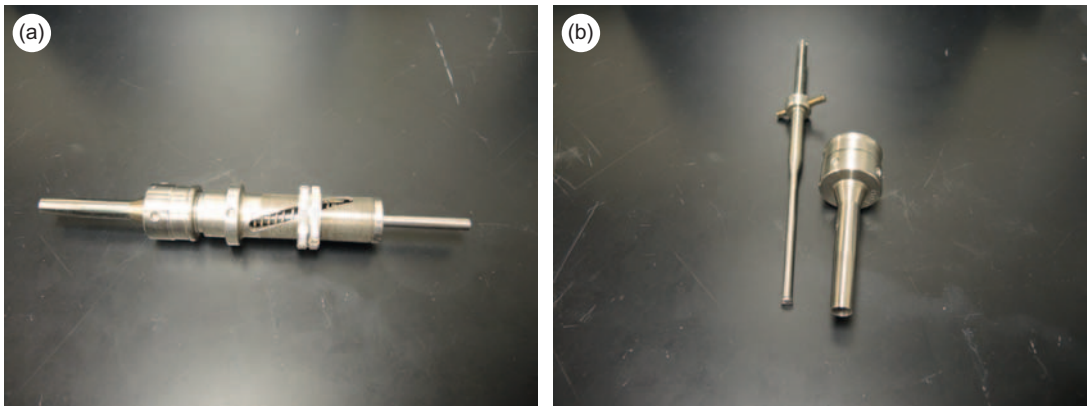


FIGURE 27.20 (a) Close up of a dosator and (b) Dosing tube and piston.

bed and positioned over an open capsule body, and the piston ejects the plug into the capsule body, (Fig. 27.21). The first step to forming a plug is to insert the dosing tube into the powder bed of a set height. This height is usually greater than the height of the piston, that is, the powder bed height is greater than the height of the open length of the tube. By inserting the dosator into a powder bed, the powder is precompressed in a manner analogous to the precompression during tablet compression in a die (Fig. 27.20). Following precompression, the powder is compressed by the piston to form a plug. Then the dosator is withdrawn from the powder bed and positioned over the empty capsule body, and the piston ejects the plug into the capsule body (Fig. 27.20). For a more detailed description of the process and the monitoring of plug formation, see Small and Augsburger.[41] Like the dosing-disk method, to have a consistent capsule fill, it is essential that the powder bed be kept at a constant bulk density. Thus, in the powder bed, there are typically moving baffles that move through the powder bed to keep the density and the height of the bed constant. This is important because the dosator is always inserted in the same location of the powder bed. The primary factor controlling fill weight is the height of the piston in the dosing tube (ie, the volume of open space in the dosing tube), and the second-most-important factor is the height of the powder bed.

27.7 CAPSULE FORMULATION REQUIREMENTS

As mentioned previously, to reliably produce capsules on a high-speed filling machine, you need at a minimum a formulation that (1) flows well, (2) consolidates to form a plug, (3) has sufficient lubricity to be efficiently ejected into the capsule body, and (4) when the capsule shell dissolves, the formulation must disintegrate for drug release. Obviously there are other criteria such as stability, but these are drug specific and beyond the scope of this chapter. In the following

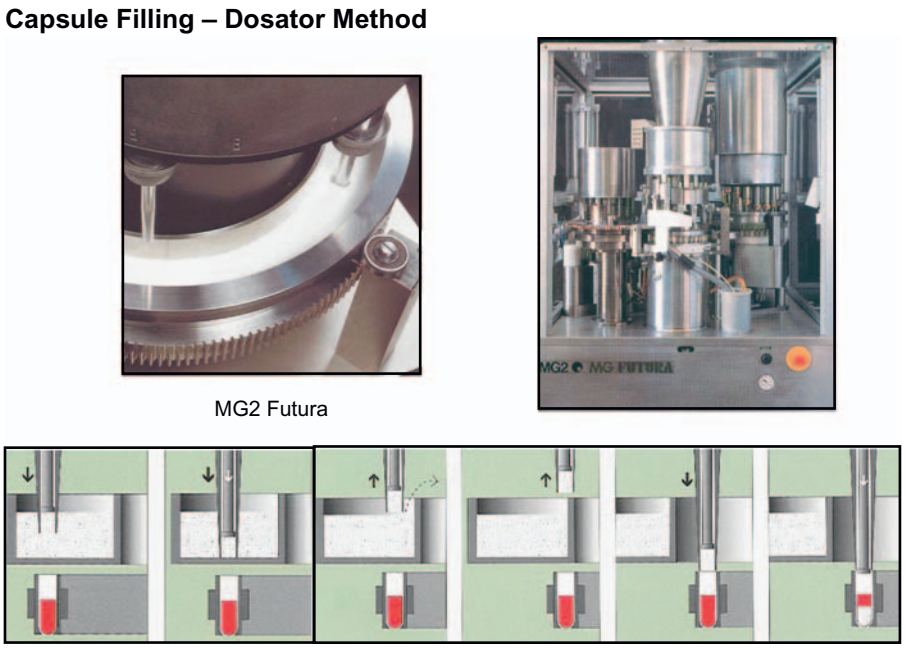


FIGURE 27.21 Dosator filling principles.

sections, we will discuss these fundamental properties that every formulation must have and the typical excipients used to create a robust formulation.

27.7.1 Flowability

For both the dosing-disk and the dosator-filling methods, a uniform powder bed density is essential to getting a consent capsule fill weight on a high-speed machine. Because for each capsule the volume of powder is repeatedly removed from the same location in the powder bed, the powder bed must reform, or the fill weight will drop and vary. Replenishing the powder bed is accomplished via the intrinsic flow properties of the powder and with the aid of baffles or agitators that move in the powder bed to redistribute the powder to the locations where the powder was removed. With all this movement of the powder bed, another formulation requirement is that the formulation does not segregate. The problem of segregation is formulation specific and beyond the scope of this chapter, but the interested reader can see Xie et al.[42] In addition, for the dosing-disk machine, fluidity is important because as the powder flows into the tamping pin holes at faster speeds, the less time the powder has to flow into the tamping pin hole, which is why sometimes with poorly flowing formulations when the speed goes up, the fill weight decreases.

As discussed by Hardy et al., flow is a critical factor affecting fill weight uniformity for dosing-disk machines.[43] Thus, when developing a formulation, the ability to measure flow in the laboratory before doing full-scale experiments on a high-speed machine will save a lot of time and resources. Much research has been done on flow measurements. One problem with most of the current flow measurements is they do not measure the fundamental physical properties, so the flow measurement can't be directly related to the performance in a capsule-filling machine. Another issue is that different flow methods look at slightly different powder flow properties. For example, in static versus dynamic flow and flow at low-bulk versus high-bulk density, the different methods won't always have the same rank correlation. Thus, one must look at several flow parameters to gain a thorough understanding of a formulation's fluidity; see Ramachandruni and Hoag for a more complete discussion of these issues.[44] Thus, with these empirical methods, experience is needed to apply them to the formulation development.

For capsule filling, some of the most commonly used methods to measure flow are the angle of repose, Carr index, Hausner ratio, funnel flow rate, FT4, and a minimum orifice diameter.[45,46] Other methods such as a shear cell are more fundamental measurements of powder flow from which flow predictions can be made, but these methods require specialized and often expansive equipment.

Two of the most commonly used methods of characterizing flow are the Carr compressibility index and the Hausner ratio. The Carr index is defined as

$$CI = \frac{\rho_t - \rho_b}{\rho_t} \times 100 \qquad (27.1)$$

where ρ_t is the tapped bulk density, and ρ_b is the loose bulk density. The Hausner ratio is defined as

$$H = \frac{\rho_t}{\rho_b} \tag{27.2}$$

Both of these methods are based upon a change in density when the powder bed is subjected to force by tapping. The principle of these methods is that when the powder is loosely packed cohesive powder will form many bridges due to the cohesive bonds between the particle at the point of contact between the particles, and when the powder bed is tapped, these cohesive bonds are broken. The powder bed collapses to a smaller volume and hence a higher density, whereas less cohesive powders won't undergo as much change in bulk density, as they do not form the bridging cohesive bonds.[48] Thus, small changes in density as a result of tapping are indicative of a good flowing powder as they don't form cohesive bonds or large changes in bulk density. From empirical experience, typical values of the Carr index and Hausner ratio for powders are given in Table 27.1. One reason these methods are so popular is that they are easy to measure, and because bulk density is important for the choice of mixer and the capsule shell size, they are often measured as part of formulation development. Because these methods are empirical, the next question is how does this correlate to fill weigh and fill weight variation, which has been the subject of much research.

Nair et al. studied the relationship between the Carr index, the flow rate index, and the minimum orifice diameter in relation to machine settings such as tamping pin height, fill weight, and fill weight variation.[49] They found that the flow is critical to the fill weight and the fill weight variation. In general, they found that pin height setting and the bed height could be set to minimize weight variation, and flow measurements were helpful in making such decisions. In another study, Osorio and Muzzio found results consistent with Nair et al., in which several flow parameters are needed to understand the relationship between weight variation and content uniformity.[50] Heda et al. compared the filling requirements of formulations run on the Zanasi LZ64 dosator and the Hofliger-Karg (H&K) GKF 70.[51] They found that poor flow measured by the Carr index was correlated by poor weight uniformity, and a Carr index in the range from 20 to 30 was ideal for weight variation.

To improve the flow of capsule formulation, glidants such as colloidal silica, cornstarch, talc, and magnesium stearate can be added to the formulation; of all these materials, colloidal silica is the most commonly used glidant. It should be noted that magnesium stearate has some effect on powder flow, but its principal effect is as a lubricant. Glidants typically have very small particle sizes and work by disrupting the cohesive forces between the particles.[52] There are many theories on how glidants function, but the mechanism by which they work is not well understood; see Armstrong for a discussion.[53] It is generally believed that to reduce cohesion between the particles, glidants coat the particles.[52] Thus, there is an optimum level for best flow, which is generally less than 1% and typically 0.1–0.25%, and empirical observation by the author has found that if the concentration of colloidal silica exceeds that needed to coat the powder surfaces, the excess colloidal silica will not improve flow and can actually reduce flowability, so the optimal glidant concentration is a critical parameter that needs to be determined.

TABLE 27.1 Empirical Relationship Between the Carr Index and Hausner Ratio

Observed flowability	Carr index	Hausner ratio
Excellent	5–10	1.05–1.11
Good	11–15	1.12–1.18
Fair	16–20	1.19–1.25
Passable	21–25	1.27–1.33
Poor	26–31	1.35–1.45
Very poor	32–37	1.47–1.59
Exceedingly poor	38–45	1.16–1.82

Adapted from Carr RL. Evaluating flow properties of solids. Chem Eng 1965; 72:415–9; Wassgren C, Pedersen H. Particle and powder flow characterization; 2010. http://pharmahub.org/resources/362.[47]

27.7.2 Compressibility and compactability of capsule plugs

With the indirect method of capsule filling, a plug is formed and then ejected into the capsule body. The typical capsule plug can be handled but will easily crumble when squeezed between your fingers. There are many similarities and important differences between capsule plug formation and tablet compaction.[43,54] Both capsule plugs and tablets are formed by the compression of a powder in a confined space. Generally, capsule plugs are compressed at forces under 200 N, while tablets are made by compression forces that range from 10 to 15 kN. A typical capsule plug has a much higher porosity, typically greater than 30%, with typical values ranging from 40% to 60%. In contrast, tablets have porosities less than 20%, and typical values are from 5% to 15%. In diametric compression, the typical breaking force for a capsule plug is under 1 N, and for tablets, the range is from 100 to 300 N. Another difference is the height-to-diameter ratio. For capsules, this is typically from 3:1 to 4:1 depending upon capsule size, and for tablets it is typically less than one, with a 0.3:1 ratio being typical.

Plug integrity is critical to the filling performance and in particular weight variation, because during the transfer from the powder bed to the capsule body, any powder that falls from the plug will result in a capsule weight variation. For dosing-disk machines, this can occur just before ejection when the dosing disk is moved over an open space where the capsule plug is unsupported to the position the plug over the capsule body. For dosator machines, this can occur when the dosing tube is withdrawn from the powder bed and moved over the capsule body.[55,56] In addition, during plug formation, differences in consolation can also lead to a plug weight variation and consequently dose uniformity problems. Thus, the ability of the plug to be transferred to the capsule body without a loss of powder is a critical quality attribute that must be controlled if a formulation is to be run successfully on a high-speed capsule-filling machine. Another key issue with plug compressibility for high-dose drugs and botanical products in particular is that you want to minimize the capsule size; thus, increasing the bulk density is important to optimizing the fill weight that can be put in the smallest size capsule possible. Note this requirement of forming a cohesive plug is at odds with the requirement of a free-flowing powder, so when developing a capsule formulation, one must balance the need for a free-flowing powder with the need to form a cohesive plug.

One way of assessing the propensity of a capsule plug to lose mass during transfer to the capsule body is to measure the mechanical strength of the plug and, in particular, plug compactability. In this chapter, I will use the notation commonly adopted by many authors such as Leuenberger: The term "compressibility" is defined as the ability of a powder to decrease in volume under pressure, and the term "compactibility" is defined as the ability of the powdered material to be compressed into a tablet of specified strength (ie, radial tensile strength or deformation hardness).[57] One of the most common ways to measure plug strength is by measuring the diametric crushing force or strength. This test is similar to measuring tablet-crushing force using diametric compression (Fig. 27.22). This method has been studied for capsule plugs.[51,55,56,58] In general, increasing the compression force increases the plug hardness and the bulk density. However, there is a point where more force, in the range of typical capsule forces, does not increase the plug-crushing force and can actually adversely affect the weight variation. The compressibility of a formulation is very dependent on the ingredients in the formulation and the compactability of the API, and each formulation must be optimized for optimal filling.

27.7.3 Lubricity

As mentioned previously, capsule plugs are typically made using compression forces less than 200 N, but forces as 300 N have been reported. While these forces are much lower than those used to make tablets, there is still a need for lubricants in a capsule formulation that is to be run on high-speed filling machines. Like tablets, traditional lubricants have been shown to significantly reduce the ejection force and improve

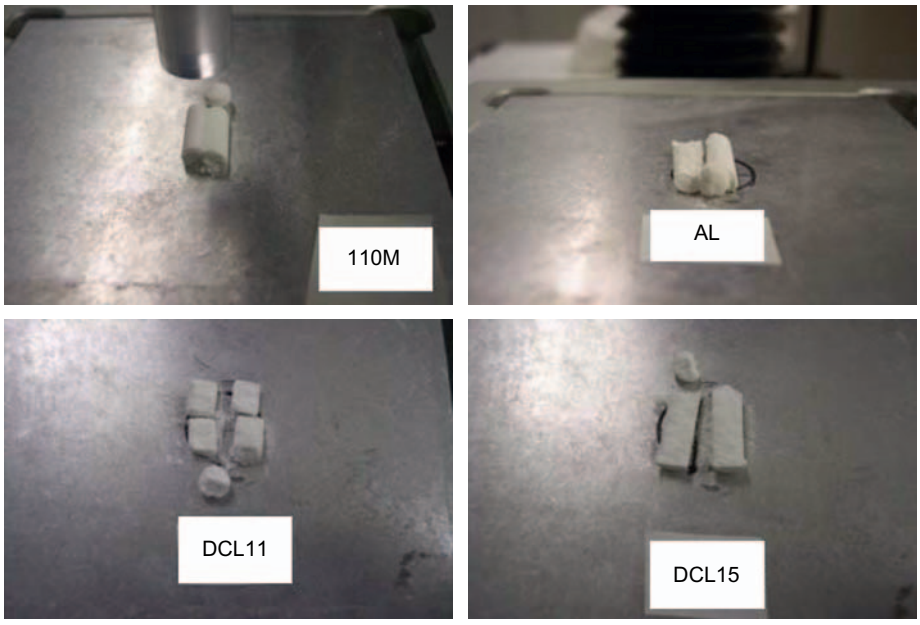


FIGURE 27.22 Diametric compress of a capsule plug; see Moolchandani et al.

capsule filling.[41,55,56] In addition, like tablets, there is an optimal lubricant level. Too little lubricant will create problems with ejection, and too much lubricant will create problems with weight uniformity and dissolution. The optimal level of lubricant depends upon the nature of the formulation. As a general rule, for plastic materials such as microcrystalline cellulose and starch, concentrations from 0.25% w/w to 0.5% w/w works best, and for brittle materials such as dibasic calcium phosphate and intermediate materials such as lactose, higher concentrations around 1.0% w/w may be needed.

27.8 CAPSULE FORMULATIONS

27.8.1 Filler binders

The nature or requirements of the capsule formulation is very dependent upon the filling method (Table 27.2). As discussed previously, for slow-speed direct-filling methods, the excipients requirements are minimal, and for PIC filling, there are no excipients in the capsule. When hand-filling using a jig, one needs a filler binder to bulk up the API volume such that the powder fill volume matches the capsule shell volume. Thus, the only excipient requirements are for a consistent bulk density, compatibility with the API, and no interference with the drug release. When filling capsules using the semiautomatic or auger principle, one needs a better flowing formulation that can undergo uniform mass flow out of a hopper without rat holing in addition to having the correct bulk density and not interfering with the drug release. When filling by the indirect methods, one needs a formulation that can form a plug, and the plug must be ejected into a capsule shell, be compatible with API, and not inhibit the drug release. To achieve these objectives, one must come up with the right combination of excipients that meets these requirements.

A typical capsule formulation for each of filling methods discussed in this chapter is shown in Table 27.2. The first thing to consider when developing a formulation is the nature of the API. The formulation must be built around the nature of the API. This discussion applies to a typical API. For very low-dose drugs of less than a couple of milligrams, special formulation and blending techniques must be applied to ensure content uniformity; see Garcia and Prescott.[59] Also, if the drug has other constraints such as low solubility, special techniques must be applied, which are beyond the scope of this chapter. If we consider the API to be one part, then the first thing to be added is the filler binder; typically the filler binder is added in two to three parts of the formulation. For example, if the dose is 50 mg, then the filler binder would typically be added in an amount of 100–150 mg as a starting point and then optimized from there.

Commonly used filler binder types include lactose, starch, dibasic calcium phosphate, and the cellulosics such as microcrystalline cellulose (MCC). Each of these excipient types comes in a variety of grades, and as a general rule for capsule filling, the direct compression grades are preferred for capsule formulation as they tend to have the best combination of flowability and compactability. When selecting the filler binder for a drug with poor solubility, it is best to choose water-soluble binders such as lactose. One should avoid high

TABLE 27.2 Typical Excipient Amounts Used in Capsule Formulations

	Direct filling methods			Indirect filling methods	
Ingredient	**PIC**	**Hand-filling**	**Auger-filling**	**Dosing disk**	**Dosator**
API	1 part	1 part	1 part	1 part	1 part
Filler-binder	NA	2–3 parts	2–3 parts	2–3 parts	2–3 parts
Disintegrants					
• Traditional • Super	NA	10%–20% wt; 4%–8% wt	10%–20% wt; 4%–8% wt	10%–20% wt; 4%–8% wt	10%–20% wt; 4%–8% wt
Lubricants					0
• Hydrophobic • Water soluble	NA	NA	0.5%–1.0% wt 2%–4% wt	0.5%–1.0% wt 2%–4% wt	0.5%–1.0% wt 2%–4% wt
Flow aids					
• Silicas • Starches	NA	NA	0.25%–0.5% wt 2%–4% wt	0.25%–0.5% wt 2%–4% wt	0.25%–0.5% wt 2%–4% wt
Surfactants	NA	0.5%–2% wt	0.5%–2% wt	0.5%–2% wt	0.5%–2% wt

concentrations of excipients that are known to gel; for example, HPMC would not be a good filler binder for low-dose drugs with poor solubility. Also, in high concentrations (ie, for low-dose drugs), MCC can form a gellike structure that can inhibit release. Thus, for a low-dose formulation, is it good to add some lactose or starch in with the MCC.

Minimizing API segregation when selecting a filler binder is a critical factor for content uniformity. As a general rule, the most important thing affecting the segregation tendency of a formulation is particle size differences between the filler binder and the API; that is, big differences in particle size should be avoided. In the author's personal experience with free-flowing noncohesive powders, ideally, I like to keep the API filler binder-to-particle size ratio somewhere between 1:3 or 3:1, with a target of 1:1. Of note, this range is the ideal range; however, in the real world, this might not be practical. Formulations with particle size ratios outside of this range may work perfectly well, but a formulator should verify the segregation tendency of the formation.[42] This also applies to free-flowing noncohesive powder systems. For cohesive powder systems, such as a micronized powder of less than 20 μm, the problem is typically not segregation but deaggregation of the micronized powder during mixing.[59]

The issue of ingredient compatibility with each other and the capsule shell is complex and very API dependent, so it won't be covered further.[60] However, the Maillard reaction between reducing sugars such as lactose glucose, galactose, maltose, and maltodextrines and drugs with primary and secondary amino groups such as peptides and proteins is a common reaction and should be considered when developing a formulation. From a practical point of view, when the reaction occurs, yellow and brown spots are formed, and as the reaction proceeds, the reactants turn black; the reaction is promoted by high temperatures and alkaline pHs.[14] Also, the Maillard reaction requires some free water, and only a small amount may be required for the reaction to occur. Thus, the hygroscopicity of the overall formulation and the type of lactose used must be considered to maintain good stability. From a moisture point of view, lactose is either in an anhydrous or a monohydrate form. The differences between them are that the monohydrate forms have a low propensity for water, while the anhydrous forms readily absorb water. In addition, the water in the monohydrate form is crystalline water and thus does not significantly react.[61,62] In summary, the capsules can be filled with lactose and have good stability for long-term storage as long as sufficient care is taken to ensure that the free water content is kept low. This requires a tight container closure system and a formulation that does not promote water absorption.[63]

27.8.2 Disintegrants

Disintegrants are added to a formulation to promote the drug release. They do this by increased water wicking into the plug, and they promote deaggregation of the plug particles. For IR tablets disintegrants are essential, but for capsules, they are less important because the plug is less of a barrier to the drug release than a compressed tablet. (See the discussion in Section 27.7.2 about the differences between a capsule plug and a tablet.) Adding a disintegrant to a capsule formulation does facilitate the drug release, which makes the formulation more rugged to the inevitable variation in excipient properties and processing conditions that occur during production over the manufacturing life of a product.

There are two classes of disintegrants: traditional disintegrants, such as starch, and super disintegrants, which include croscarmellose sodium, crospovidone, and sodium starch glycolate. Currently, these three supper disintegrants are the most popular disintegrants. In capsule formulations, super disintegrants are typically used at levels from 4% to 8% (Table 27.2), which is about twice that used in a typical tablet formulation.[64,65] The reason for this is that when a disintegrate breaks apart the plug, it does so by swelling, and this expansion of the disintegrant particles pushes the adjacent particles apart, but because plugs have a much higher porosity and the particles are not packed as close together, a given amount of expansion does not push the particle apart as much as in a tablet where the particles are much closer together. Sodium starch glycolate is known to swell more than other super disintegrants, which makes it a popular choice for capsule formulations.

27.8.3 Lubricants and flow aids

The requirements for formulation lubricity and flowability have been discussed, and in this section, the lubricants and flow aids for capsule formulation will be discussed. In common usage, the terms lubricant and flow aid are sometimes used synonymously, but for this discussion, we will use the term lubricant or true lubricant to describe an excipient that reduces the friction between the capsule plug and the metal surface of the dosing tube or the dosing disk. A glidant is something that improves the flowability of the powder. In addition, sometimes lubricants are added to reduce powder adhesion to metal surfaces. For some drugs, such as amlodipine, adherence to metal surfaces can be a significant problem; if the powder builds up, it can adversely affect content uniformity.

Lubricants can be broken into hydrophobic and water soluble. Of the hydrophobic lubricants, the most popular and efficient lubricant is magnesium stearate. In this context, efficiency refers to the ability

of a lubricant to reduce powder metal friction. Hydrophobic lubricants include the metallic stearates such as calcium stearate and stearic acid. Magnesium stearate typically uses about 0.5% wt, while the metallic stearates are included at between 0.5% and 1.0% wt, and stearic acid in the range of 1–5% wt is commonly used. For some capsule formulations, it is important that all the ingredients be soluble; for example, capsules used in a diagnostic kit need to be completely soluble. For these formulations, lubricants such as polyethylene glycol (PEG) can be used. These lubricants are not as efficient as magnesium stearate and must be used at a higher concentration; in the range of 5–15% wt is typical. To improve flow, the most common glidant is colloidal silica, but also starch and talc can be used to improve flow in a capsule formulation; the typical amounts used are given in Table 27.2.

27.8.4 Surfactants

For some drugs with low bioavailability, surfactants can be added to increase the wetting of powder mass. By improving API wetting, this can increase the rate of water uptake, which can improve the dissolution and the bioavailability of the formulation. Two commonly used surfactants are sodium docusate (SDS) and sodium lauryl sulfate (SLS). Typical use levels for SDS are in the range of 0.1–0.5% wt, and for SLS, they are in the range of 1–2% wt (Table 27.2).

References

1. Murachanian D. Capsule shells and the marketing of capsules. In: Augsburger LL, Hoag SW, editors. *Pharmaceutical dosage forms: capsules*. London: Taylor Francis; 2016.
2. Edwards D. Applications of capsule dosing techniques for use in dry powder inhalers. *Ther Deliv* 2010;**1**:195–201.
3. Overgaard ABA, Moller-Sonnergaard J, Christrup LL, Hojsted J, Hansen R. Patients' evaluation of shape, size and colour of solid dosage forms. *Pharm World Sci* 2001;**23**:185–8.
4. Abdul S, Chandewar AV, Jaiswal SB. A flexible technology for modified-release drugs: multiple-unit pellet system (MUPS). *J Control Release* 2010;**147**:2–16.
5. Stegemann S. Non-gelatin based capsule. In: Augsburger LL, Hoag SW, editors. *Pharmaceutical dosage forms: capsules*. London: Taylor Francis; 2016.
6. Duconseille A, Astruc T, Quintana N, Meersman F, Sante-Lhoutellier V. Gelatin structure and composition linked to hard capsule dissolution: a review. *Food Hydrocol* 2015;**43**:360–76.
7. Jones BE, Podczeck F, Lukas P. Capsule shell manufacture. In: Augsburger LL, Hoag SW, editors. *Pharmaceutical dosage forms: capsules*. London: Taylor Francis; 2016.
8. Coppola M, Djabourov M, Ferrand M. Unified phase diagram of gelatin films plasticized by hydrogen bonded liquids. *Polymer (Guildf)* 2012;**53**:1483–93.
9. Djabourov M, Grillon Y, Leblond J. The sol-gel transition in gelatin viewed by diffusing colloidal probes. *Polymer Gels Networks* 1995;**3**:407–28.
10. Jones RT. Gelatin: manufacture and physico-chemical properties. In: Podczeck F, Jones BE, editors. *Pharmaceutical capsules*. London: Pharmaceutical Press; 2004. p. 23–60.
11. GMIA GMIOA. 2012. *Gelatin handbook*. www.gelatin-gmia.com/images/GMIA_Gelatin_Manual_2012.pdf.
12. Digenis GA, Gold TB, Shah VP. Cross-linking of gelatin capsules and its relevance to their in vitro-in vivo performance. *J Pharm Sci* 1994;**83**:915–21.
13. Ofner CM, Zhang Y-E, Jobeck VC, Bowman BJ. Crosslinking studies in gelatin capsules treated with formaldehyde and in capsules exposed to elevated temperature and humidity. *J Pharm Sci* 2001;**90**:79–88.
14. Wu Y, Levons J, Narang A, Raghavan K, Rao V. Reactive Impurities in Excipients: Profiling, Identification and Mitigation of Drug–Excipient Incompatibility. *AAPS PharmSciTech* 2011;**12**:1248–63.
15. Bowtle W, Kanyowa L, Mackenzie M, Higgins P. Physical stability and resistance to peroxidation of a range of liquid-fill hard gelatin capsule products on extreme long-term storage. *Drug Dev Ind Pharm* 2011;**37**:685–93.
16. Meyer MC, Straughn AB, Mhatre RM, Hussain A, Shah VP, et al. The effect of gelatin cross-linking on the bioequivalence of hard and soft gelatin acetaminophen capsules. *Pharm Res* 2000;**17**:962–6.
17. Jones BE. Gelatin alternatives and additives. In: Podczeck F, Jones BE, editors. *Pharmaceutical capsules*. London: Pharmaceutical Press; 2004. p. 61–77.
18. Nyamweya N, Hoag SW. Influence of pigments in the properties of polymeric coating systems. In: Felton L, McGinity J, editors. *In Aqueous polymeric coatings for pharmaceutical dosage forms*. (3rd Edition). New York, NY: Informa Healthcare; 2016. p. 151–76.
19. Jones BE. Manufacture and properties of two-piece hard capsules. In: Podczeck F, Jones BE, editors. *Pharmaceutical capsules*. London: Pharmaceutical Press; 2004. p. 79–100.
20. Augsburger LL. Hard and soft shell capsules. In: Banker GS, Rhodes CT, editors. *Modern pharmaceutics*. New York, NY: Marcel dekker, Inc; 1996. p. 395–440.
21. Barham AS, Tewes F, Healy AM. Moisture diffusion and permeability characteristics of hydroxypropylmethylcellulose and hard gelatin capsules. *Int J Pharm* 2015;**478**:796–803.
22. Silva SMC, Pinto FV, Antunes FE, Miguel MG, Sousa JJS, Pais AACC. Aggregation and gelation in hydroxypropylmethyl cellulose aqueous solutions. *J Colloid Interface Sci* 2008;**327**:333–40.
23. Joshi SC. Sol-Gel behavior of hydroxypropyl methylcellulose (HPMC) in ionic media including drug release. *Materials* 2011;**4**:1861.
24. Funami T, Hiroe M, Noda S, Asai I, Ikeda S, Nishinari K. Influence of molecular structure imaged with atomic force microscopy on the rheological behavior of carrageenan aqueous systems in the presence or absence of cations. *Food Hydrocol* 2007;**21**:617–29.
25. Morris ER, Rees DA, Robinson G. Cation-specific aggregation of carrageenan helices: domain model of polymer gel structure. *J Mol Biol* 1980;**138**:349–62.
26. Moritaka H, Nishinari K, Taki M, Fukuba H. Effects of pH, potassium chloride, and sodium chloride on the thermal and rheological properties of gellan gum gels. *J Agric Food Chem* 1995;**43**:1685–9.
27. Ku MS, Lu Q, Li W, Chen Y. Performance qualification of a new hypromellose capsule: part II. Disintegration and dissolution comparison between two types of hypromellose capsules. *Int J Pharm* 2011;**416**:16–24.
28. Sherry Ku M, Li W, Dulin W, Donahue F, Cade D, et al. Performance qualification of a new hypromellose capsule: Part I. Comparative evaluation of physical, mechanical and

processability quality attributes of Vcaps Plus®, Quali-V® and gelatin capsules. *Int J Pharm* 2010;**386**:30–41.
29. Cole ET, Scott RA, Cade D, Connor AL, Wilding IR. In vitro and in vivo pharmacoscintigraphic evaluation of ibuprofen hypromellose and gelatin capsules. *Pharm Res* 2004;**21**:793–8.
30. Khawam A. Modeling powder encapsulation in dosator-based machines: I. Theory. *Int J Pharm* 2011;**421**:203–9.
31. Khawam A, Schultz L. Modeling powder encapsulation in dosator-based machines: II. Experimental evaluation. *Int J Pharm* 2011;**421**:210–19.
32. Adeyeye MC, Adebayo A. Hard shell capsules in clinical trials. In: Hoag SW, Augsburger LL, editors. *Pharmaceutical dosage forms: capsules*. London: Taylor Francis; 2016.
33. Rahman MA, Hussain A, Hussain MS, Mirza MA, Iqbal Z. Role of excipients in successful development of self-emulsifying/microemulsifying drug delivery system (SEDDS/SMEDDS). *Drug Dev Ind Pharm* 2013;**39**:1–19.
34. Cole ET, Cadé D, Benameur H. Challenges and opportunities in the encapsulation of liquid and semi-solid formulations into capsules for oral administration. *Adv Drug Deliv Rev* 2008;**60**:747–56.
35. Stegemann S, Richardson M. Filling two-piece hard gelatin capsules with liquids. *Tablets Capsules* 2007;**5**(2):81–6.
36. Ay. Tamper-resistant packaging requirements for certain over-the-counter (OTC) human drug products. *Federal Register (USA)* 1989;**54**:5227–9.
37. Niederquell A, Kuentz M. Introduction of a theoretical splashing degree to assess the performance of low-viscosity oils in filling of capsules. *AAPS PharmSciTech* 2011;**12**:323–30.
38. Hoag SW, Dave V, Moolchandani V. Compression and compaction. In: Hoag SW, Augsburger LL, editors. *Pharmaceutical dosage forms: tablets*, vol. 1. New York, NY: Informa Healthcare; 2008. p. 555–630.
39. Lightfoot D. Hard shell capsule filling machines. In: Hoag SW, Augsburger LL, editors. *Pharmacueutical dosage forms: capsules*; 2016. London: Taylor Francis.
40. Shah K, Augsburger LL, Small LE, Polli GP. Instrumentation of a dosing disk automatic capsule filling machine. *Pharm Technol* 1983;**7**:42–54.
41. Small LE, Augsburger LL. Instrumentation of an automatic capsule filling machine. *J Pharm Sci* 1977;**66**:504–9.
42. Xie L, Wu H, Shen M, Augsburger LL, Lyon RC, et al. Quality-by-design (QbD): effects of testing parameters and formulation variables on the segregation tendency of pharmaceutical powder measured by the ASTM D 6940-04 segregation tester. *J Pharm Sci* 2008;**97**:4485–97.
43. Hardy IJ, Fitzpatrick S, Booth SW. Rational design of powder formulations for tamp filling processes. *J Pharm Pharmacol* 2003;**55**:1593–9.
44. Ramachandruni H, Hoag SW. Design and validation of an annular shear cell for pharmaceutical powder testing. *J Pharm Sci* 2001;**90**:531–40.
45. Carr RL. Evaluating flow properties of solids. *Chem Eng* 1965;**72**:415–19.
46. Hao T. Understanding empirical powder flowability criteria scaled by Hausner ratio or Carr index with the analogous viscosity concept. *RSC Adv* 2015;**5**:57212–15.
47. Wassgren C, Pedersen H. 2010. Particle and powder flow characterization. http://pharmahub.org/resources/362.
48. Wu C-Y, Armstrong B, Vlachos N. Characterization of Powder Flowability for Die Filling. *Particulate Science & Technology* 2012;**30**:378–89.
49. Nair R, Vemuri M, Agrawala P, Kim S-i. Investigation of various factors affecting encapsulation on the In-Cap automatic capsule-filling machine. *AAPS PharmSciTech* 2004;**5**:46–53.
50. Osorio JG, Muzzio FJ. Effects of powder flow properties on capsule filling weight uniformity. *Drug Dev Ind Pharm* 2013;**39**:1464–75.
51. Heda PK, Muteba K, Augsburger LL. Comparison of the formulation requirements of dosator and dosing disc automatic capsule filling machines. *AAPS Pharmsci* 2002;**4**:E17.
52. Augsburger LL, Shangraw RF. Effect of glidants in tableting. *J Pharm Sci* 1966;**55**:418–23.
53. Armstrong NA. Lubricants, glidants, and antiadherents. In: Hoag SW, Augsburger LL, editors. *In pharmaceutical dosage forms: tablets*, vol. 2. New York, NY: Informa Healthcare; 2008. p. 251–67.
54. Jones BE. The filling of powders into two-piece hard capsules. *Int J Pharm* 2001;**227**:5–26.
55. Moolchandani V, Augsburger LL, Gupta A, Khan M, Langridge J, Hoag SW. Characterization and selection of suitable grades of lactose as functional fillers for capsule filling: part 1. *Drug Dev Ind Pharm* 2015;**41**:1452–63.
56. Moolchandani V, Augsburger LL, Gupta A, Khan MA, Langridge J, Hoag SW. To investigate the influence of machine operating variables on formulations derived from lactose types in capsule filling: part 2. *Drug Dev Ind Pharm* 2016;**42**:624–35.
57. Leuenberger H. The compressibility and compactibility of powder systems. *Int J Pharm* 1982;**12**:41–55.
58. Heda PK, Muller FX, Augsburger LL. Capsule filling machine simulation. I. Low-force powder compression physics relevant to plug formation. *Pharm Dev Technol* 1999;**4**:209–19.
59. Garcia TP, Prescott JK. Blending and blend uniformity. In: Augsburger LL, Hoag SW, editors. *Pharmaceutical dosage forms: tablets*, vol. 1. New York, NY: Informa; 2008. p. 111–74.
60. D'Hondt M, Wynendaele E, Vandercruyssen K, Bauters T, Vandenbroucke J, et al. Investigation of active pharmaceutical ingredient loss in pharmaceutical compounding of capsules. *J Pharm Biomed Anal* 2014;**96**:68–76.
61. Du JP, Hoag SW. The influence of excipients on the stability of the moisture sensitive drugs aspirin and niacinamide: comparison of tablets containing lactose monohydrate with tablets containing anhydrous lactose. *Pharm Dev Technol* 2001;**6**:159–66.
62. Wang T, Alston KM, Wassgren CR, Mockus L, Catlin AC, et al. The creation of an excipient properties database to support quality by design (QbD) formulation development. *Am Pharm Rev* 2013;**16**:16–25.
63. Pharma D. Lactose: some basic properties and characteristics. www.dfepharma.com/en.
64. Botzolakis JE, Augsburger LL. Disintegrating agents in hard gelatin capsules .1. Mechanism of action. *Drug Dev Ind Pharm* 1988;**14**:29–41.
65. Botzolakis JE, Augsburger LL. Disintegrating agents in hard gelatin capsules .2. Swelling efficiency. *Drug Dev Ind Pharm* 1988;**14**:1235–48.

This page intentionally left blank

CHAPTER

28

Design, Development, and Scale-Up of the High-Shear Wet Granulation Process

S. Badawy and P. Pandey

Bristol-Myers Squibb, New Brunswick, NJ, United States

28.1 INTRODUCTION

The high-shear wet granulation process is one of the more traditional and commonly used granulation processes in design and development of a solid oral drug product. It is used to achieve the desired bulk density, flow, and compaction properties of the formulation.[1] It also helps minimize segregation and offers better content uniformity.[2] In certain instances it can also be used to improve the bioavailability of a formulation by providing intimate contact between a poorly water-soluble drug and a surfactant, and was shown to mitigate food effect.[3,4] Another known benefit of the wet granulation process is the robustness of the process to handle changes in active pharmaceutical ingredient (API) powder properties. This can be important during the early stages of drug product development where the API manufacturing process undergoes optimization and scale-up, thus producing materials with some variations in powder properties. Other manufacturing processes, such as direct compression and dry granulation, usually experience a more significant impact of any changes to API powder properties.[5,6] Direct compression has the advantage of short manufacturing time, but requires formulation with adequate density, flow, and compaction properties. These criteria are not commonly met for high drug loading formulations. On the other hand, low drug loading formulations may present a high risk of blend segregation and poor content uniformity when using a direct compression process. There are also some significant challenges with dry granulation (using roller compaction), such as flow into the roller compactor (especially at high drug loading), sticking to the rolls, and loss of compaction after passing through the rolls,[7,8] all of which are not present for the high-shear wet granulation process. Wet granulation is the preferred manufacturing method for formulations with high drug loading. However, the use of water and/or high degree of shear during wet granulation offer a different set of challenges for APIs that are sensitive to water and shear and can undergo transformation (eg, form conversion) that may lead to a worse drug product stability profile.[9–14] The wet granulation process also requires an additional unit operation of drying. There are other variants of wet granulation process that may be used as alternatives to the high-shear process, such as low-shear, fluid bed, and twin-screw granulation based on the specific formulation and process challenges of the drug product. For APIs sensitive to shear, a low-shear process may be preferrable.[15] Fluid bed granulation has the advantage of providing short exposure of the granulation to the water compared with high-shear granulation. It imparts less shear and densification of the granulation and results in more porous granules that can have a higher dissolution rate and better compaction properties. A choice between these manufacturing processes can be made by manufacturing process risk assessments based on prior drug product formulation knowledge, such as sensitivity to moisture, shear, and/or temperature.

A typical high-shear wet granulation formulation consists of API, filler (eg, microcrystalline cellulose and/or lactose), disintegrant (eg, superdisintegrants such as croscarmellose sodium), binder (eg, polyvinylpyrrolidone (PVP), hydroxypropyl cellulose (HPC)), and lubricant (eg, magnesium stearate). In some cases it may also contain a surfactant (eg, sodium

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00028-5

lauryl sulphate) and a flow aid (eg, silicon dioxide). The components that are added in the granulation bowl and undergo granulation are called intragranular components and are incorporated into the granules. The use of a polymeric binder in this step helps to form solid bridges between the particles when the granules are dried up, which dictates mechanical properties of the formed granules in conjunction with process parameters. The extragranular components of a typical formulation can include the use of filler (for providing compaction enhancement), disintegrant (to help disintegrate the overall tablet), and lubricant (aids in compaction process). Extragranular components are generally kept low in concentration (unless needed for a specific purpose) to reduce the possibility of blend segregation during blend transfer and compaction process.

A high-shear wet granulation process involves a granulation bowl equipped with an impeller blade that provides agitation and imparts high-shear to the powder during granulation. The bowl contains a chopper that helps to break-up large agglomerates formed during granulation. The binder is usually added (sprayed through a nozzle or pumped through a tube) as part of a binder solution from the top of the bowl. The binder may also be added in a dry manner, in which it is included as part of the powder formulation and water alone is used as a granulating liquid. This mode of dry binder addition is gaining popularity because it removes the additional processing steps of ensuring binder solution mixing, dispersion, and appropriate storage.[16,17] However, it can be argued that dry binder addition does not allow enough time for the binder to be fully hydrated, and thereby renders it less effective.

A schematic of a high-shear wet granulator is shown in Fig. 28.1.[18] The overall process can be subdivided into various steps. The first step ("preblending" or dry mixing) involves charging the intragranular portion of the formulation into the granulator bowl and mixing all the components to ensure a uniform mixture. The second step is when the granulating liquid is added at a certain rate to the moving powder bed, and is called the "water-addition" step. In this stage the binder gets distributed on the powder surfaces as the powder is moved by the impeller, promoting particle-particle collisions and thereby causing particle agglomeration. In the third and final step of granulation, the wet mass of granules is mixed in the granulator for some additional time, and is termed the "wet-massing" step. This step ensures further uniform distribution of water and is mainly dominated by granule consolidation and breakage.

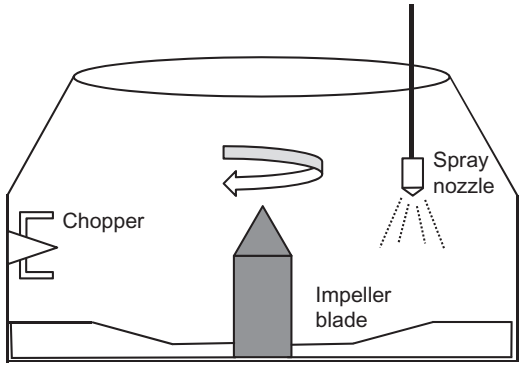


FIGURE 28.1 Schematic of a high-shear wet granulation process.

At the end of granulation, the wet granules typically require a wet milling step through a relatively coarse mesh to break up oversized agglomerates that may have resulted from local overwetting. Routinely at the end of granulation, a normal distribution of granule size is not obtained and there is a portion of large-size agglomerates that require some form of milling. At a small-scale (<200 g), this can be done by passing the wet granules through a coarse mesh (eg, US mesh size 8). For pilot and larger scale batches, a mill may be used to eliminate the coarser granules and obtain a more uniform granule size distribution. The coarser wet granules (if not eliminated) carry an additional risk during the drying step, especially when drying in a fluidized bed, where the larger granules will fluidize differently than the rest of the granules and may not dry as efficiently. Large granules have a higher risk of "case hardening," where granules dry mostly on the outside surface while retaining a wet core.

After the wet milling, the following step is to dry these granules, which is done either using a tray oven or a fluidized bed. The choice between the drying equipment is often based primarily on the batch size, but other factors such as the physical/chemical characteristics of materials, production capacity, drying characteristics, and initial and final desired moisture content all play a role in the selection. Tray oven is often used for early-stage development work (eg, formulation screening studies) where batch sizes are small (<500 g). During tray drying, the granules should be spread over a larger area using multiple trays so as not to form a deep granule bed, or else caking of powder might be observed during drying. This can then have an impact on the granule particle size distribution (and sometimes granule porosity), especially if a high amount of binder is present in the formulation.

Fluidized bed drying is the method of choice for pilot-scale and commercial scale batches. Fluidized bed drying is a more efficient and quicker process than the tray drying method and produces more

porous granules, but relies on the ability of the powder to be fluidized appropriately. Too high of an airflow at the start of the process may result in a wet mass of granules being carried all the way up the fluidized bed, sticking to the filter bags, and this wet mass may not be dried completely and remain stuck to the filter bags. On the other extreme, too low of airflow will result in a "bubbling" fluidized bed that does not provide efficient drying. The risks of case hardening and powder consolidations exist if the process conditions are not chosen appropriately (eg, high temperatures and low airflow/fluidization). Fluidized bed drying method is likely to result in more fines than the tray drying method due to higher granule attrition in the fluid bed. A comparison of granule size distribution before and after the fluidized bed drying process can give an indication of the granule strength and contributes to deciding the amount of binder to be used in the formulation. Granules coming out of the dryer are passed through a screen (typically using a conical mill) that helps to break-up any large granules and result in a more uniform granule size distribution, thus reducing segregation potential. The dry milled granules are blended with the extragranular excipients and lubricant in the final blending step(s) to form the "final blend," which is then compressed into tablets.

28.2 RATE PROCESSES IN WET GRANULATION

The high-shear wet granulation is a complex process with several and simultaneous rate processes, such as nucleation, coalescence, consolidation, and attrition, governing the overall granulation outcome. The net effect of all the rate processes depends on the processing conditions and also on the formulation components (material properties). For example, higher impeller speed may promote higher aggregation, but may also promote higher attrition, depending on the other process and formulation factors. This section provides a fundamental description of the different factors governing each of the rate processes.

28.2.1 Liquid distribution and nucleation

As liquid is added to the granulator, it wets the powder particle and creates liquid bridges between them, resulting in the formation of initial granules or nuclei. Wetting of the powder particle by the liquid is essential for successful nucleation and is governed by the spreading coefficient of the liquid on the solid, λ_{LS}:

$$\lambda_{LS} = W_A - W_{CL} \tag{28.1}$$

where W_A is the work of solid-liquid adhesion and W_{CL} is the work of cohesion of the liquid.[19,20] A positive spreading coefficient indicates that spreading of the liquid on the solid surface is thermodynamically favorable and conducive to nucleation.

In addition to the thermodynamics of wetting, size distribution of the formed nuclei is also dependent on kinetics of nucleation.[21] Kinetics of nucleation is a function of liquid addition and powder mixing parameters. In pharmaceutical wet granulation, liquid is traditionally added as bulk by pouring into the granulator bowl, metered though tubing using a pump, or atomized with a nozzle. Granule size distribution is dependent on droplet size of the liquid in cases where liquid is added by atomization into small droplets with an addition rate that is low relative to the speed of powder surface and droplet penetration time into the powder bed. If this is the case, powder flux in the spray zone is fast and droplet penetration into the bed is rapid compared with liquid addition rate and thus, liquid droplets do not overlap on the surface of the powder bed. Nuclei size distribution is determined by droplet size distribution of the granulating liquid.

Dimensionless spray flux[22–27] is a measure of the density of droplets falling on the powder bed surface, and is defined by Eq. (28.2):

$$\Psi_a = \frac{\dot{a}}{\dot{A}} = \frac{3\dot{V}}{2\dot{A}d_d} \tag{28.2}$$

where Ψ_a is the dimensionless spray flux, $\dot{A}$ is the powder flux or spray area in the nucleation zone per unit time, $\dot{V}$ is the volumetric spray flow rate, d_d is the average droplet size, and $\dot{a}$ is the total projected area of drops per unit time. A very low value of Ψ_a ($\ll 1$) indicates low droplet density, which do not coalesce with one another before hitting the powder surface.

Droplet penetration time into a powder bed, t_p, is given by:

$$t_p = 1.35\frac{V_d^{2/3}}{\varepsilon^2}\frac{\mu}{R_{eff}\gamma_{lv}\cos\theta} \tag{28.3}$$

which shows that droplet penetration time is increased (slower penetration) at higher liquid viscosity, μ, and decreased at higher bed porosity, ε, and bed pore radius, R_{eff}.[28,29] V_d is drop volume, γ_{lv} is liquid surface tension, and θ is liquid-solid contact angle.

Hapgood et al.[25] proposed a nucleation regime map in terms of spray flux and drop penetration time (made dimensionless by normalizing to particle

circulation time). In regions with low spray flux (<0.1) and fast drop penetration, each individual droplet forms a separate nucleus granule, and the nucleation process is described as being in the "droplet-controlled" regime. In this case, the initial nuclei size will be directly related to the droplet size, and therefore can be controlled by atomization parameters. On the other hand, a very high value of Ψ_a (>1) will result in the spray area being covered by droplets that have coalesced and the individual initial droplet size information bears little meaning. In that case, the water distribution is governed by the mechanics of the granulator (eg, impeller speed and design, chopper speed and design) and is called the "mechanical dispersion" regime. Local overwetting may occur when operating in the mechanical dispersion regime if shear forces are not sufficient for effective liquid distribution. This was shown to result in initial bimodal distribution of the granulation, which slowly changes to unimodal distribution by continued mixing in the granulator.[30] Nucleation is always in the mechanical dispersion regime when liquid is added as bulk by pouring into the granulator.

28.2.2 Consolidation

Porosity of granules decreases by mixing in the high-shear granulator as granules collide with each other, with the granulator wall, and with the impeller. Forces imparted on the granules as the result of such collisions cause primary particles within the granules to move closer to each other, resulting in reduced granule porosity. As porosity is decreased, granule strength increases, which makes it more resistant to breakage. Interparticle frictional forces are typically the main barrier to consolidation of pharmaceutical granules, particularly for cohesive materials with small and wide particle size distribution.[31,32] Granulating liquid acts as a lubricant which reduces interparticle friction and thus facilitates granule consolidation.[33]

28.2.3 Coalescence and growth

As granules collide in the granulator they can either stick together (coalesce), resulting in growth, or they may bounce back with no increase in granule size. Layering is the term used for a special type of coalescence when fine particles or granules coalesce with larger ones. Probability of successful coalescence is inversely related to particle size and as granule size in the granulator increases, rate of coalescence decreases and granule growth tends to slow down. Coalescence is a dynamic process that begins as soon as nuclei are formed during the initial phase of liquid addition and continues until a maximum granule size is achieved (a_{cr}^{coal}). The probability of successful coalescence between granules depends on mechanical properties of the granules and on the presence of granulating liquid on the granule surface. Probability of coalescence is enhanced by the increase in plastic deforming ability of granules because granule deformation increases the surface area of granule contact available for bonding between granules. Granule deformation also squeezes liquid to the granule surface, which forms liquid bridges between colliding granules and aids in coalescence.[34]

Granules liquid saturation, which is defined as the portion of pores within a granule that are filled with liquid, was shown to be a key factor in controlling granule coalescence.[34] Granule liquid saturation takes into account both the volume of available pores as well as the liquid volume as follows:

$$S = \frac{w\rho_s(1-\varepsilon)}{\rho_l \varepsilon} \tag{28.4}$$

Where S is granule liquid saturation, w is the mass ratio of liquid-to-solid, ρ_s is solid density of primary particles, ρ_l is liquid density, and ε is granule porosity. As liquid saturation increases, granules become more plastic and the thickness of liquid layer on granule surface increases. Granule growth was found to increase at higher liquid saturation because coalescence is enhanced by both of these factors.[34–36] Liquid saturation increases as granules consolidate and porosity decreases, and thus, consolidation was found to enhance granule growth in many cases.

Iveson and Litster[37] proposed a granule growth regime map in which granule growth behavior is dependent on granule deformation behavior (defined by a granule deformation number) and granule pore saturation. They described two regions of granule growth behavior, steady growth and induction growth, depending on granule deformability and pore saturation. Steady growth occurs in deformable granules and is characterized by linear increase in granule size as function of time. Induction growth is exhibited by strong granules with little deformation upon impact resulting in an initial phase of minimal growth. Once sufficient consolidation occurs, liquid saturation is increased which is followed by a phase of rapid growth.[38] In addition to steady growth and induction growth regions, the regime map also included a nucleation-only region at low liquid saturations, where nuclei can form but coalescence does not take place due to the low liquid saturation. To the contrary, at very high liquid saturations, rapid growth and eventually slurry formation can occur. For very weak materials, a crumb region is described where permanent granules cannot form.

Different models have been reported for granule coalescence, which can be broadly classified into two classes.[21] In Class I models,[39,40] coalescence takes place if the kinetic energy of colliding granules is fully dissipated during collision. As a result, granules do not bounce back and it is assumed that the formed bond between the granules is sufficiently strong to result in permanent coalescence. Class II models[31,41] assume that granules are sufficiently plastic to dissipate the kinetic energy and establish contact between the granules for a finite amount of time. However, depending on the strength of the formed bond, granules may permanently stick together or separate due to shear forces in the granulator. The model of Liu et al.[39] is a prominent example of Class I models. They proposed two types of coalescence for deformable surface wet granules. In type I coalescence, kinetic energy of colliding granules is dissipated by the viscous forces in the liquid layer on granule surface before their surfaces come in contact with each other. This results in the formation of a larger granule consisting of the two colliding granules held together by a liquid bridge. Type I coalescence can also occur for nondeformable granules because it does not involve granule deformation. In type II coalescence, kinetic energy is not completely dissipated by the viscous forces in the liquid layer. Further dissipation of the kinetic energy occurs via plastic deformation as granule surfaces come in contact, resulting in complete loss of the kinetic energy. For nonsurface wet deformable granules, only type II coalescence is possible.

28.2.4 Attrition and breakage

High-shear mixers typically have a wide distribution of particle velocities and shear forces. A granule formed by coalescence in one region of the granulator can break up as it goes through another region with a higher shear intensity. Granule breakage was shown by tracer studies using colored tracer granules.[42,43] Tardos et al.[44] defined a limiting granule size (a_{cr}^{break}) above which a granule will break, which is inversely related to shear rate. This is analogous to the limiting granule size for successful coalescence, a_{cr}^{coal}, which was mentioned previously. Granules would grow up to the a_{cr}^{coal}, but would tend to break if a_{cr}^{coal} exceeds a_{cr}^{break}. Large granules formed by local overwetting in the liquid addition zone will also tend to break up if they are larger than a_{cr}^{break}. As mixing continues in the granulator, a steady state with a stable granule size distribution may be reached that depends on the relative magnitude of $a_{cr}^{break}$and a_{cr}^{coal} and distribution of shear forces in the granulator.

TABLE 28.1 A Summary of Potential CMAs, CPPs and Potential Critical Granule Properties as they Relate to High-Shear Wet Granulation

Potential CMAs (Input material properties)	Potential CPPs (Process parameters)	Potential critical granule properties
Particle size distribution	Impeller speed	Particle size distribution
Surface area	Wet massing time	Pore volume distribution
Particle shape and morphology	Water addition time	Granule morphology
Solubility	Spray rate	Granule mechanical properties and strength
Binder viscosity	Nozzle properties	Bulk, tap density
Surface tension of granulating liquid	Water amount (water to solids ratio)	Flowability
Bulk density	Bowl fill	Compaction properties
Water uptake capacity	Granulator geometry	Content uniformity
Contact angle (wettability)	Impeller design/ configuration	
Particle surface properties		

Source: From Pandey P, Badawy SI. A quality by design approach to scale-up of high-shear wet granulation process. Drug Dev Ind Pharm 2016;42(2):175-189

It should be obvious from the previously stated overview that high-shear wet granulation is a complex unit operation that involves multiple simultaneous rate processes. In order to design a robust drug product and achieve the desired target product profile (TPP), it is important to have a good understanding of the critical material attributes (CMAs) and critical process parameters (CPPs) of the process and how these CMAs and CPPs are related to the drug product critical quality attributes (CQAs). A summary of potential CMAs, CPPs, and potential critical granule properties impacted by them is shown in Table 28.1, and are discussed in detail in the subsequent sections.

28.3 MATERIAL PROPERTIES IN WET GRANULATION

There are various physical attributes of the excipients and API that govern the granulation process and thereby the final characteristics of the resultant drug product. While it may be thought that physical properties of the API are more important than the

"inactive excipients," the latter play an equally important role in the granulation process. Examples of physical properties of input materials (excipients and drug substance) that have been shown to have an effect on granulation process include particle size and shape, surface area, solubility, and contact angle with the binder solution. The granulation phenomenon is also affected by the properties of the granules that are formed during the process, such as the granule strength and granule porosity. A discussion on some of the main attributes is provided in this section.

28.3.1 Powder properties

28.3.1.1 Particle size

Strength of wet agglomerates is directly impacted by particle size distribution of the starting materials. According to the model proposed by Rumpf,[45] the static strength of moist agglomerates of a single component is expressed as:

$$\sigma_T = SC\left[\frac{(1-\varepsilon)}{\varepsilon}\right]\left[\frac{\lambda_v \cos\theta}{d_p}\right] \tag{28.5}$$

where σ_T is the tensile strength of a moist agglomerate, S is liquid saturation of the agglomerate, ε is agglomerate porosity, λ_v is liquid surface tension, θ is liquid-solid contact angle, d_p is diameter of primary particles (assuming monosized spheres), and C is a material constant. While the Rumpf equation was derived for a single component with monosized particles, it can be possibly useful for a mixture of components with a particle size distribution using relevant approximation of mean particle size and liquid-solid contact angle.

Small particle size of the starting material also results in moist agglomerates that are more resistant to densification and produce more porous granules under similar process conditions.[33] Moist compacts prepared with smaller particle size lactose were more resistant to densification in a uniaxial compaction test than the larger particle size fraction, which was also in agreement with the higher porosity of the granules prepared in the high-shear granulator using the material with smaller particle size.[46] Kristensen et al. also reported a similar observation for dicalcium phosphate.[31] They proposed a model relating agglomerate tensile strength to intrinsic interaction parameter of particles within the agglomerate and showed that this interaction parameter increased with the decrease in particle size and increase in the width of particle size distribution.[32] Granule porosity for the development compound DPC 963 was also found to be higher when using smaller drug substance particle size under the same process conditions[47] (Fig. 28.2).

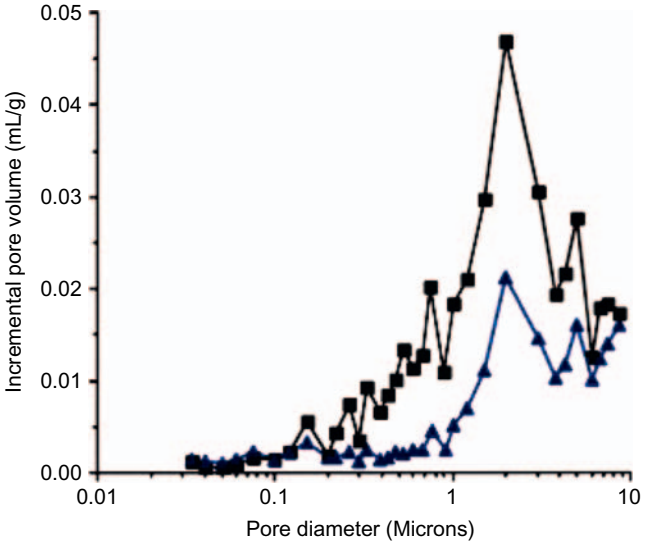


FIGURE 28.2 Porosity of DPC 963 granulation fraction manufactured using small (■) and large (▲) drug substance particle size. *Source: From Badawy SI, Lee TJ, Menning MM. Effect of drug substance particle size on the characteristics of granulation manufactured in a high-shear mixer.* AAPS PharmSciTech *2000;1(4):E33.*

The increase in granule strength at the smaller particle size is attributed to the increased volume density of interparticle contacts, which consequently increases frictional forces that contribute to the strength of wet granules.[21] The higher strength of moist agglomerates at the liquid distribution stage is conducive to more robust nucleation and reduced attrition. Increasing material particle size decreases strength of formed granules and moves the systems up the growth regime map described in Section 28.2.3 from induction to steady growth and eventually to the crumb region. A minimum particle size is hence required for a given solid/liquid system above which stable granules cannot form.[21] At the coalescence phase, granule growth is enhanced by higher plasticity and lower yield stress of the granules.[39] Therefore, at the same porosity and liquid saturation, granules formed from larger particle size material are expected to have lower yield stress and more rapid coalescence. However, the effect of particle size on coalescence is confounded by its effect on consolidation tendency and, thus, liquid saturation. Granules with larger starting material particle size are more easily densified and possess lower porosity and higher liquid saturation at the same solid-to-liquid ratio. The lower porosity increases yield stress while the higher liquid saturation tends to increase plasticity and reduces yield stress.[46] The effect of starting material particle size on yield stress and granule coalescence will thus depend on the system-dependent balance of these opposing mechanisms. For many

pharmaceutical systems, increased porosity outweighs the impact of lower liquid saturation[46] and therefore smaller starting material particle size was found to enhance granule growth in many cases.[29,46,47] In their study of 25 pharmaceutical compounds and excipients, Vemavarapu et al. found a statistically significant enhancement in granule growth resulting from the decrease in starting material particle size.[29] However, an opposite effect of particle size on granule growth was reported for dicalcium phosphate, where larger particle size showed faster growth.[31] Dicalcium phosphate demonstrated induction growth with brittle granule behavior and minimal coalescence until very high liquid saturations. The impact of increased liquid saturation for the low-porosity granules with larger stating material particle size appeared to be predominant in this case. For cohesive materials like dicalcium phosphate, the effects of increased liquid saturation and diminished interparticle friction are significant; resulting in lower yield strength and rapid coalescence for the denser agglomerates prepared using the larger particle size material.

28.3.1.2 Surface area

Surface area determined by a gas adsorption method represents the sum of external and internal (within particle) surface area. Surface area is inversely related to particle size for materials with minimal internal porosity or surface irregularities and therefore increased surface area resulting from the decrease in particle size of these materials is expected to have the same effects as described above for particle size. However, the effect of a material's internal surface area on the wet granulation process is mediated through a different mechanism and should therefore be considered independently from particle size. As a result, Vemavarapu et al.[29] did not find a correlation between the effect of inverse particle size and surface area on granule growth. Granule growth was found to decrease as both surface area and particle size increased. Inclusion of both terms in the regression model was necessary to optimize model fit. It appears, therefore, that the two factors affect granule growth through independent mechanisms. The contribution of internal porosity to surface area was more significant than the impact of particle size on surface area. Surface area was primarily construed as the internal area (intraparticle porosity) with additional contribution of surface irregularities. The penetration of water into the pores of the primary particles during granulation makes less of it available at the particle surface which decreases liquid film thickness and granule pore saturation, thus reducing the chances of coalescence and growth.

28.3.1.3 Contact angle

Wetting of the powder particles by the granulating liquid (low solid-liquid contact angle) is necessary for robust nucleation. High contact angle results in slow nucleation and low strength of wet granules. Slow nucleation is the result of decreased spreading coefficient and the less thermodynamically favorable wetting at the higher contact angle. In addition, the static strength of the wet granule is also lowered at high contact angles, as predicted by the Rumpf equation (Eq. (28.5)). Once formed, the wet granule should withstand the high-shear and impact forces generated in the mixer. Such lack of strength due to poor wetting results in breakage of wet granules, thus reducing granule growth. For example, granulation of sulfur with different polymer solutions was not successful due to the poor wetting of the sulfur particles by the granulating liquid.[35] The increase in solid-liquid contact angle by the addition of the hydrophobic salicylic acid to lactose resulted in the reduction of the resulting granule particle size.[21] Vemavarapu et al.[29] reported low growth ratio (defined as ratio of granulated blend particle size to particle size of starting material) at the same granulation conditions for materials with a high contact angle with water. Jaiyeoba and Spring also reported a similar negative influence of high contact angle and poor wettability on granule growth.[21,48]

Poor wettability and high contact angle also results in decreased tendency for densification. Poor wetting by the granulating liquid reduces its effectiveness in lubricating and reducing fractional forces at interparticle contact points, which resists movements of particles.[37] Consequently, poor wetting of the particles by the granulating liquid results in porous, weaker granules.

28.3.1.4 Solubility

During wet granulation, dissolution of soluble components in the granulating liquid takes place as liquid is added and mixed with the solids in the granulator. If the formulation contains a large fraction of a highly soluble active or excipient, a significant portion of the formulation can dissolve during wet granulation, resulting in the decrease in the solid-to-liquid ratio and an increase in granulating liquid viscosity. The decrease in solid-to-liquid ratio shifts the system to the rapid growth and overwet mass regions of the growth regime map. The increased viscosity can also affect granule growth as described next. Consequently, formulations with highly soluble component(s) require a lower amount of granulating liquid to achieve desired growth and to avoid "overgranulation" and uncontrolled granule growth. Granule growth was

reported to be uncontrolled in highly soluble compositions, as reflected in the large oversize fraction of the resulting dried granules (defined as weight fraction of granules >1.4 mm).[29] In the same study, highly soluble compositions also demonstrated wider size distribution for the dried granules, which was attributed to the higher tendency for uncontrolled growth. The increased viscosity of the granulating liquid upon dissolution of the highly soluble material, which hinders liquid distribution, was also suggested as a contributing mechanism to the formation of the wide granule size distribution.

28.3.2 Granulating liquid properties

28.3.2.1 Viscosity

Granulating liquid viscosity impacts all rate processes in wet granulation: nucleation, consolidation, coalescence and breakage.[11] Since its effect on the different rate processes can be in opposite directions, the impact of viscosity on final product attributes is system dependent. Oftentimes, the binder (eg, PVP or HPC) is dissolved in the granulating liquid, which increases its viscosity depending on binder type, concentration, and molecular weight. High viscosity impedes liquid dispersion for systems in the mechanical dispersion nucleation regime, resulting in localized overwetting. Longer times and higher impeller energy are required for uniform dispersion of the viscous granulating liquid and, subsequently, achieving uniform growth takes a longer time. Short or low-energy granulation processes with viscous granulating liquid often result in oversize granules and wide granule size distribution. Granulation of calcium carbonate with PEG 4000 aqueous solutions resulted in a wide bimodal granule size distribution attributed to the poor dispersion of the highly viscous binder.[49] For systems operating in the droplet-controlled nucleation regime, the increase in binder viscosity may lead to larger droplet size and increased droplet penetration time, which can eventually shift the system to the mechanical dispersion regime. Dry mixing of the binder with other formulation ingredients and granulating with water is a frequently used strategy to reduce viscosity and improve granulating liquid dispersion. However, rapid hydration of the binder as water is added to the formulation may still impede rapid dispersion of the granulating liquid.

High viscosity increases strength of the liquid bridges between primary particles within the granule according to Eq. (28.6)[11]:

$$F_v = \frac{3\pi\mu r_p^2}{2h}\frac{dh}{dt} \tag{28.6}$$

where F_v is the dynamic strength of the liquid bridge, μ is liquid viscosity, r_p is the primary particle radius, h is the distance between the two primary particles held by the liquid bridge, and t is time. The increased strength of liquid bridges resists particle movement within the granule, thus decreasing consolidation and resulting in more porous granules.[40] In addition, higher viscosity hinders liquid movement through the intragranular space, which makes liquid squeezing during consolidation more difficult.[33] Accordingly, the strength of dried calcium carbonate granules was reduced by the increase in polyethylene glycol 4000 concentration and viscosity of the granulating liquid. The higher viscosity of the more concentrated PEG 4000 solutions decreased granule consolidation propensity resulting in more porous and weaker final granules.[49] Similarly, Iveson et al. showed that the rate of granule consolidation increases as liquid viscosity decreases.[33]

Granulating liquid viscosity has a pronounced effect on the different coalescence mechanisms described in Section 28.2.3 (type I and type II with surface wet or surface dry granules). Higher binder viscosity on granule surface dissipates kinetic energy of colliding granules and increases probability of successful type I coalescence for surface wet granules.[40] Viscous Stokes number (St_v), which describes the balance between kinetic energy of colliding granules and dissipating viscous forces in the liquid bridge, decreases with the increase in liquid viscosity as follows:

$$St_v = \frac{8\rho_g r u}{9\mu} \tag{28.7}$$

where ρ_g is granule density, r is the mean granule radius, u is relative velocity of granule impact, and μ is the viscosity of the liquid on granule surface. A limiting value for St_v exists for each system below which successful type I coalescence takes place.

For deformable granules, the higher liquid viscosity and enhanced granule strength increases yield stress, which increases the Stokes deformation number (St_{def}) and decreases probability of type II coalescence. St_{def}, a measure of the ratio of granule kinetic energy to energy lost by plastic deformation of colliding granules, is described as[39]:

$$St_{def} = \frac{mu^2}{2D^3 Y_g} \tag{28.8}$$

where D and m are the mean granule diameter and mass, respectively, and Y_g is the granule yield stress. For surface wet deformable granules, increased binder viscosity has two competing effects on coalescence with the net effect determined by the contribution

of each of them to coalescence in a given system. For surface dry deformable granules, higher binder viscosity can only reduce coalescence.

28.3.2.2 *Surface tension*

Static strength of moist agglomerates is enhanced by the higher surface tension of the liquid according to the Rumpf equation (Eq. (28.5)). Increased strength of the moist agglomerates at the nucleation phase reduces their attrition, resulting in more robust nucleation. Maximum stable granule size is increased with higher surface tension granulating liquids.

The stronger liquid bridges formed by granulating liquids with higher surface tension are also more capable of pulling primary particles together, resulting in lower granule porosity. Granule density increased progressively with the reduction in surface tension as ethanol was added to water in the granulating liquid.[50] Ritala et al. also found that granule porosity was lower when the higher surface tension PVP solution was used to granulate dicalcium phosphate compared with the lower surface tension PVP/polyvinyl alcohol copolymer and hydroxypropyl methylcellulose (HPMC) solutions.[35] Growth was also faster with the higher surface tension PVP solutions, which was attributed to the lower porosity and enhanced liquid saturation. The authors found that granule growth is strongly dependent on liquid saturation and showed that relationship between granule size and liquid saturation is independent of the granulating solution used.

The lower strength of granules due to low liquid surface tension makes plastic granules more easily deformable, which increases probability of type II coalescence. However, the low surface tension decreases the bond formed between the two colliding granules, which makes the resulting granule more susceptible to breakage.[11]

28.4 DESIGN OF THE PHARMACEUTICAL WET GRANULATION PROCESS

There is a significant interaction between formulation components and process variables that dictate the granulation outcome. The two main components of the process include the granulator equipment and granulation liquid. The factors associated with granulator equipment would be the granulator design, bowl volume and fill, impeller blade design (size, shape, angle), impeller blade placement (top or bottom driven), impeller and chopper speed, chopper design, and process time. The factors associated with granulation liquid include amount, liquid introduction method (spray vs drip), spray rate, spray characteristics (droplet size, velocity, spray area, etc.), liquid addition time, addition location, liquid properties (viscosity, surface tension, etc.), and binder addition method (wet vs dry). The previous section discussed the input material properties (formulation aspects), while this section focuses on the main process variables that govern the wet granulation process.

28.4.1 Impeller and chopper speeds

Speed and frequency of particle collisions in the granulator increase as the impeller speed is increased. Increasing collision energy enhances deformability and growth of plastic particles,[39] while increasing attrition in brittle granules with low liquid saturation.[51] At low water level and liquid saturation, granules are more brittle and high collision velocity is more likely to result in granule breakage rather than coalescence and growth. In other words, granule behavior when yield stress is exceeded determines the effect of impeller speed. For brittle granules, which break at their yield stress, a maximum stable granule size is achieved under given conditions according to the model proposed by Tardos et al.,[44] which predicts the maximum stable granule size to be inversely related to impeller speed.

$$a_{\mathrm{cr}} = \frac{(2\tau_{\mathrm{y}} St_{\mathrm{def}}/\rho_p)^{1/2}}{\gamma} \tag{28.9}$$

where a_{cr} is the is the maximum stable granule size, τ_{y} is the yield stress, ρ_{p} is the granule density, γ is the average shear rate in the granulator, and St_{def} is the Stokes deformation number. This model predicts the maximum stable granule size to be inversely related to impeller speed because γ is directly related to the impeller speed.

In contrast to brittle granules, plastic granules, which deform rather than break at their yield stress, demonstrate increase in probability of coalescence with the increase in impeller speed, according to the model of Liu et al.[39] The effect of impeller speed on granule size thus depends on granule mechanical properties where plastic granules are likely to have more pronounced growth as the impeller speed is increased, while brittle granules can undergo more attrition and reduction in granule size at the higher impeller speed. Consequently, the increase in granule size with impeller speed was found to be dependent on water level used for granulation in case of razaxaban and brivanib formulations.[51] For both formulations, the increase in granule size at the higher impeller speed was predominantly observed at the high water level where granules have sufficient liquid saturation and plasticity.[51] Other

studies also reported increased granule growth rate at higher impeller speed.[52,53]

In addition to the granule mechanical properties, the effect of impeller speed on granule size can also depend on the impeller's effect on liquid distribution in the granulator. As pharmaceutical high-shear wet granulation usually operates in the mechanical dispersion regime for nucleation where mechanical forces are responsible for liquid distribution and nucleation,[25] a low impeller speed that does not efficiently distribute the liquid can result is localized overwetting and formation of lumps.[51] These lumps or oversize granules are distinct, at least from a mechanistic point of view, from large granules that result from granule coalescence and uniform growth. The oversize lumps represent a larger particle size than would be achieved if the granulating liquid is uniformly dispersed and the granulation process is allowed to proceed to steady state. Higher impeller speed can thus increase the efficiency of water distribution and reduces oversize lumps and granule size is also reduced as a result. Consequently, pexacerfont formulation showed significant reduction in granule size as the impeller speed was increased.[51] In this case, the reduction in particle size at the high impeller speed was more predominant at the higher water level. This was attributed to the improved water distribution at the higher impeller speed and the breakage of lumps resulting from localized overwetting at the high water level.[51]

The higher impeller speed also results in higher consolidation forces in the granulator and decreased granule porosity.[49,51,52,54] As the porosity is decreased, granule liquid saturation is increased at the same level of granulating liquid. According to Iveson and Litster's regime growth map,[37] this increase in liquid saturation shifts the system's behavior from the nucleation to the growth regions, resulting in faster granule growth as the increased consolidation squeezes the liquid to the granule surface, which enhances coalescence.

Increasing impeller speed may result in a more spherical shape for deformable granules at high liquid saturation.[54,55] Granule shape can affect its performance as more spherical granules have higher bulk density due to their closer packing ability. Spherical granules will thus tend to have improved flow properties. On the other hand, more irregularly shaped granules may have improved volume reduction behavior and compactability.[55]

Chopper speed is not usually as critical as impeller speed. Chopper may be useful in breaking up overwetted lumps and aids in liquid distribution. Chopper may also provide shear forces contributing to granule growth. In addition, the chopper may disrupt uniform flow pattern of powders and improve mixing in the granulator. Knight[53] reported narrowing of the granule size distribution when the chopper was switched on, which was attributed to the chopper's preferential breakage of large granules. Despite the narrower distribution, mean granule size distribution was not impacted by chopper action.

28.4.2 Amount of granulating liquid

The ratio of liquid to solid used for granulation determines granule liquid saturation, which researchers reported to be a critical factor with respect to granule growth and consolidation.[32,35,36,51,52,56–59] As granule liquid saturation is increased, granules move from the nucleation to the growth regions on the growth map. Increased liquid saturation makes granules more deformable and enhances coalescence.[33] Granulating liquid also lubricates particle movements within the granules, reduces interparticle friction, and facilitates consolidation. Consequently, granule porosity is reduced as liquid saturation in increased. Iveson et al.[33] reported decreased granule porosity at the higher liquid saturation when low viscosity liquid was used for granulation. An opposite effect was found when high viscosity liquid was used. In the latter case, the disruption of interparticle friction by the liquid was outweighed by the increased strength of the liquid bridges of the viscous liquid, which resists densification.

28.4.3 Process duration

Process duration is usually defined as the time taken from the beginning of granulating liquid addition until mixing is stopped. It is typically broken down into two phases: (1) liquid addition phase during which granulating liquid is added to the granulator while mixing, and (2) wet massing time which is the term used to describe mixing that takes place in the granulator after granulating liquid addition is complete. It is usually beneficial to delineate the duration of the two phases as their impact on the granulation process may be different, and thus, controlling the duration of one of the phases may be more effective than controlling total process duration.

Wet massing phase helps provide even distribution of water along with some further agglomeration, reduction of fines, consolidation, and attrition of granules. During wet massing, granule coalescence and growth may take place, but large granules may also undergo breakage until a steady state particle size distribution is achieved.[44] Wet massing can result in an increase[53,60] or decrease[61] in particle size. Increasing wet massing time has been reported to reduce granule d_{90} due to attrition of large granules, while increasing

d_{10} and d_{50} resulting in a narrower granule size distribution.[58] Wet massing time in pharmaceutical wet granulation, however, is not usually long enough to achieve true steady state.

During wet massing, granules that were formed during water addition can also get further consolidation. Granule densification, which takes place during wet massing, can affect granule liquid saturation and mechanical properties and thus, particle size distribution of the granules. For systems with induction behavior, consolidation, and subsequent increase in liquid saturation during wet massing is required to induce coalescence and granule growth. Typically, an exponential decay of porosity is observed during this phase[62,63] as shown in Eq. (28.10).

$$\frac{\varepsilon-\varepsilon_{\min}}{\varepsilon_0-\varepsilon_{\min}}=\exp(-k_c t) \tag{28.10}$$

where ε is the average granule porosity, $\varepsilon_{\min}$ is the minimum achievable porosity at given process conditions, ε_0 is the initial porosity, k_c is the granulation consolidation rate constant and t is the wet massing time. Consolidation rate during wet massing is usually also dependent on other process conditions, including impeller speed and the amount of water used during granulation.

The wet massing time should generally be kept short because higher porosity is generally a desired property. High-shear wet granulation processes should therefore be designed with sufficient granulating liquid added to achieve adequate granule liquid saturations to be in the steady growth regime[37] at the end of liquid addition phase, which results in coalescence and desired granule growth. Design of induction systems that require long wet massing times and consolidation to increase liquid saturation and induce growth should be avoided. Increasing the granulating liquid amount to achieve desired growth is thus, a preferred strategy than using lower liquid amount and long wet massing time. Despite the short wet massing used in steady growth systems, significant reduction in porosity and subsequent decrease in compactability of granules and dissolution rate of tablets have been reported.[64] Increasing the wet massing time for brivanib from 10 to 50 s showed a statistically significant decrease in porosity and tablet dissolution rate. Significant increase in granule density and reduction in compactability was observed for razaxaban upon the increase in wet massing time from 30 to 180 s. Similarly, reduction in granule porosity and compactability was reported for a lactose-based granulation[52] (Fig. 28.3) and for a lactose/microcrystalline cellulose system[65] upon increasing wet massing time to 90 and 180 s, respectively. The high liquid saturation for the steady growth systems makes them prone to rapid densification and therefore design strategy should aim to minimize or eliminate[52] wet massing time.

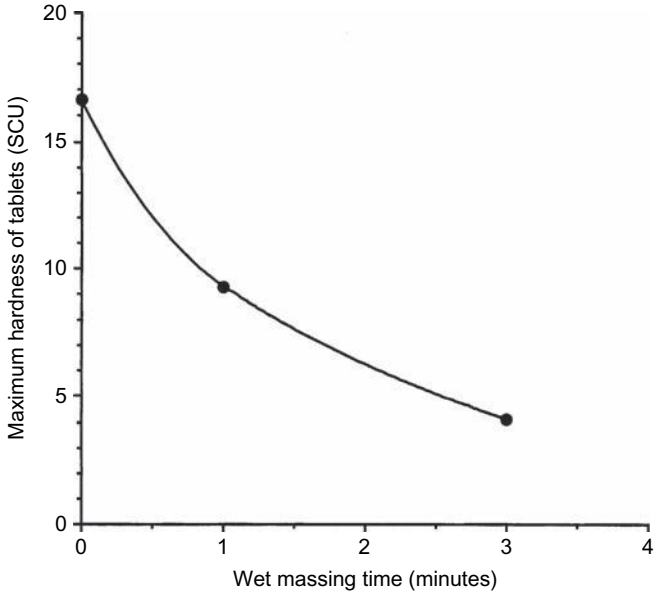


FIGURE 28.3 Effect of wet massing time on compressibility of granulation manufactured in the PMA-65 high-shear granulator. *Source: From Badawy SI, Menning MM, Gorko MA, Gilbert DL. Effect of process parameters on compressibility of granulation manufactured in a high-shear mixer.* Int J Pharm *2000;**198**(1):51–61.*

While the increase in mixing time after complete water addition (wet massing time) resulted in a significant effect on granule densification, the increase in mixing time during water addition (water addition time) showed minimal impact on granule porosity for the same formulations.[52,64] This suggests that the effect of mixing on granule porosity is dependent on the water level (granule liquid saturation). Mixing after complete water addition tends to be more effective in decreasing granule porosity (due to the lubricant effect of water mentioned previously). On the other hand, mixing during water addition, when granule liquid saturation is still low, is not as effective in reducing granule porosity and consequently prolonging water addition time has little or no effect on granule porosity. It is therefore prudent to control wet massing time as a separate parameter distinct from water addition time and not to lump both as one parameter representing total processing time.

In many cases, liquid addition rate mainly affects liquid dispersion under conditions where mechanical dispersion of the granulating liquid is not very effective. In these instances, reducing liquid addition rate aids in the dispersion of the liquid, decreases overwetted lumps, and produces narrower granule size distribution. This is more likely to be the case for viscous granulating liquids, when the binder is dissolved in the granulating liquid, or if the formulation

contains a high concentration of polymeric material such as cellulosic, polymer-based, controlled-release matrix formulation that hydrates rapidly when it comes in contact with water. Water addition time had a statistically significant effect on mean granule size for pexacerfont and the increase in water addition time (slower water addition) resulted in smaller granule size.[51] A statistically significant term for the two-way interaction between water amount and water addition time was also observed, which showed a more pronounced effect of water addition time at the higher water level. Batches manufactured at the higher water level are more likely to have a higher content of the overwetted lumps and thus, the slower water addition is more beneficial in improving water dispersion for those batches.[51] Particle size of paracetamol granules manufactured at the low impeller speed and highest liquid addition rate showed the widest distribution, likely because of the poor liquid distribution at these conditions.[66]

28.5 QUALITY ATTRIBUTES OF WET GRANULATED PRODUCTS

Microscale granule properties, such as size, morphology, porosity, and strength dictate the macroscale attributes of the granulation and final product. In the previous sections, the impact of starting material attributes and process parameters on granule microstructure were discussed. In this section, the effect of wet granulation process on performance-indicating macroscale product quality attributes is reviewed.

28.5.1 Solid state form

Solid state phase transformation is a common occurrence in wet granulation. Phase transformation in wet granulation can be classified into the following categories: (1) transformation to a more stable polymorphic form, (2) formation of a hydrate/solvate, (3) formation of a high-energy metastable form (eg, amorphous), and (4) transformation of a salt form to the less-soluble free form. Solid state transformations present quality and regulatory challenges to drug product development and commercial manufacturing. Thus, careful risk assessment, mechanistic understanding, and appropriate control strategy should be in place if wet granulation is to be used for a drug substance that is prone to such transformations.

Transformation of a metastable polymorph to the more thermodynamically stable polymorph can take place via solvent-mediated mechanism during wet granulation. For example, transformation of form A to the more stable tautomeric polymorph B was reported to take place during wet granulation of an irbesartan formulation.[67] Polymorphic conversion of the metastable form I of flufenamic acid to the stable polymorph III was also observed during wet granulation with ethanol.[68] As the more soluble metastable polymorph dissolves in the granulating liquid, it forms a supersaturated solution with respect to the stable polymorph, which subsequently crystallizes from the supersaturated solution. The less-soluble polymorph may have lower dissolution rate and bioavailability compared with the metastable form.[69,70] The extent of conversion to the more stable polymorphic form depends on the kinetics of phase transformation relative to processing time. Rate of phase transformation is a function of dissolution rate of the metastable form, volume of granulating liquid used in the process, and crystallization kinetics of the stable polymorph.[71] Extent of transformation can thus be dependent on granulation parameters that impact these factors. Processing time is dependent on liquid addition rate and wet massing time, while dissolution rate may be a function of impeller speed.

Compounds which exist in a hydrate form can potentially convert to the hydrate if the anhydrous form is used in aqueous wet granulation because water activity during wet granulation is usually higher than the critical humidity for the formation of the hydrate. Hydrate formation can also take place through a solvent-mediated mechanism, as described previously for polymorphic transformations. Many literature articles exist on the transformation of the anhydrous form to hydrate during wet granulation.[11–13] The lower solubility of hydrates can potentially lead to a lower dissolution rate and bioavailability.[70,72] The hydrate form may partially or completely convert back to the anhydrous form during drying of the granulation, so a mixture of forms can potentially exist in the final product. Wikström et al.[12,13] used Raman spectroscopy to follow hydrate formation of theophylline during wet granulation. They concluded that hydrate formation takes place via a solvent-mediated mechanism and showed that the rate of transformation increased by the increase in agitation speed. Formation of theophylline monohydrate in the wet mass was inhibited by granulation at elevated temperature (50°C).[73] Theophylline monohydrate, formed by granulation at lower temperatures (27–40°C), was found to convert back to the anhydrous form by drying at 70°C.[73]

Hydrate formation of an excipient can also take place during wet granulation. Formation of lactose monohydrate upon wet granulation of anhydrous lactose was reported.[74] Anhydrous lactose has superior compaction properties compared with lactose monohydrate due to its particle morphology.[74] Anhydrous lactose particles

have high surface roughness and porosity, in contrast with the smooth low-porosity lactose monohydrate particles. Interestingly, lactose monohydrate particles created from anhydrous lactose particles in wet granulation still retained the morphological features of the initial anhydrous lactose particles and consequently maintained higher compactability compared to the "as is" lactose monohydrate.[74]

Formation of a high-energy metastable form can take place when the starting lower energy polymorph dissolves in the granulating liquid and is kinetically trapped into a high energy form during drying.[71] A high energy disordered phase can also form through a nonsolvent mediated mechanism where shear forces applied during wet granulation in the presence of water cause disruption of crystallinity and create crystal defects or amorphous phase. Formation of an amorphous phase can also increase formulation hygroscopicity after wet granulation.[75] Wet granulation of BMS-561388 resulted in partial transformation to the less stable, amorphous form.[9] Similarly, aqueous wet granulation of either polymorph A or B of dexketoprofen trometamol with microcrystalline cellulose caused transformation to the amorphous form regardless of the initial polymorph used in granulation.[76]

Salts of weak bases or weak acids may covert to the free form (free base or free acid) during aqueous wet granulation. As the salt dissolves in the granulating water, it will form a saturated solution with respect to the free form if the solution pH is higher (for a weak base) or lower (for a weak acid) than the pH of maximum solubility (pH_{max}). The free form subsequently precipitates from this saturated solution, resulting in the salt-to-free-form transformation. If this solution pH requirement is not met, the salt would be the thermodynamically stable form in equilibrium with solution and consequently there would be no conversion to the free form. As for polymorphic transformations, wet granulation parameters (amount of granulating liquid, impeller speed, and process duration) can potentially affect the extent of solvent-mediated transformation of the salt to the free form.

Excipients can be used to stabilize a drug substance against form conversion during wet granulation. Excipients may inhibit solvent-mediated polymorphic and hydrate transformations by absorption of the solvent (water), thus making it less available to dissolve the active. In addition, specific interactions between a drug molecule and an excipient can slow down nucleation and/or crystal growth. In a study of the effect of several polymers on hydrate formation of three model compounds in wet granulation, Gift et al.[11] found varying degrees of inhibitory effect on hydrate formation depending on the compound and polymer used. Cross-linked poly(acrylic) acid was found to completely inhibit caffeine hydrate transformation and both HPMC and cross-linked poly(acrylic) acid completely inhibited the carabamazepine hydrate formation. On the other hand, sulfaguanidine transformation was rapid even in the presence of polymers. Airaksinen et al.[77] found that hygroscopic, partially crystalline excipients hindered hydrate formation of nitrofurantoin in wet granulation, likely due to their significant uptake of granulating water. Form transition from a metastable to a stable form of theophylline during wet granulation was also inhibited by povidone. Higher molecular weight of povidone was more effective in inhibiting the transition than the lower molecular weight polymer.[78]

28.5.2 Chemical stability

The wet granulation process was reported to affect chemical stability of drug molecules. In some cases, the effect of wet granulation process on chemical stability manifests itself by enhancing degradation reactions during processing whereby an increase in degradant content is observed at the end of manufacturing. Exposure of a moisture-sensitive compound to the high water activity experienced in aqueous wet granulation may be the underlying mechanism. Alternatively, degradation in the solution phase, as a fraction of the active dissolves in the granulating liquid, can be the causal mechanism. Actives with high solubility in the granulating liquid are expected to be more susceptible to the latter mechanism. Degradation during manufacturing, however, is not frequently observed in wet granulation due to the short processing time relative to reaction rates even for moisture-sensitive compounds.[9,10,79]

Enhanced degradation caused by wet granulation is more commonly observed upon storage (accelerated and long-term stability) and not at the end of manufacturing. While this may be attributed in some cases to a high level of residual water in the drug product, it is more frequently the result of creation of high-energy disordered states during wet granulation. Rate of degradation reaction in the high-energy phase is significantly higher than the crystalline phase and thus drug degradation in the solid state usually takes place in disordered regions or in crystal defects.[80] The higher molecular mobility in those disordered regions results in a faster degradation rate than in the crystal lattice. In addition, disordered regions have higher water content due to absorbed water, which further enhances degradation rate in those less-ordered regions for reactions in which water acts as a reactant.[81] As described in the previous section,

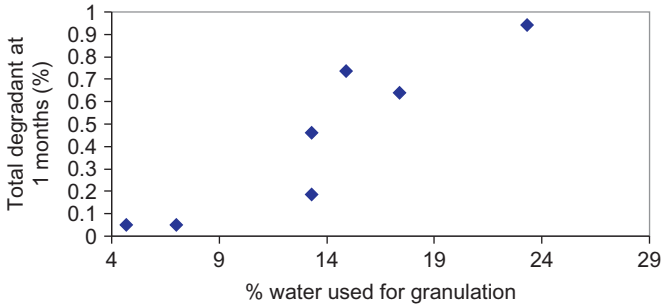


FIGURE 28.4 Effect of water amount used for granulation on the stability of compound I tablets at 40°C/75% RH.

amorphous regions of the active could form in wet granulation, and although the fraction of the drug in amorphous state may be small, this still leads to a measurable impact on stability because the acceptable specification limit of a drug product degradant is usually very low.[81] The rate of degradation of a highly water-soluble triazine derivative (compound I) on accelerated stability was thus dependent on the amount of water used for granulation (Fig. 28.4), despite the similar granulation moisture content at the end of drying. This was attributed to the higher amorphous content resulting from the increased fraction of the active dissolving in the larger volume of granulating liquid. The fraction dissolving during granulation was shown to form amorphous material after drying.[10]

In some cases, wet granulation can result in the formation of a high-energy state of one or more excipients, which increases formulation hygroscopicity and negatively impacts stability of moisture-sensitive compounds. For example, wet granulation increased moisture uptake and degradation rate of the unbuffered lactose-based formulation of the low dose compound, DMP 754, which was attributed to the formation of amorphous lactose during wet granulation.[75]

The effect of wet granulation on product stability can also be caused by enhanced contact between drug and excipients. Partial dissolution of an excipient and/or the drug substance during wet granulation and subsequent crystallization during drying (albeit to the same initial form) results in higher surface area of drug/excipient contact. This intimate contact of the drug and the excipient enhances rate of drug degradation caused by drug-excipient interaction and results in a less stable drug product.

The intimate contact of drug and excipients resulting from wet granulation was leveraged to enhance stability in some cases. For formulations containing a pH modifier, dissolution of the buffer components in the granulating liquid results in better distribution of the pH modifier in the formulation and in a more effective pH control, thus maximizing drug product stability.[82] Thus, while the rate of acid-catalyzed hydrolysis of compound I was enhanced by wet granulation compared with a dry blend in formulations without pH modifier, the opposite effect of wet granulation was observed in formulations containing sodium carbonate as a pH modifier.[10] The wet granulation process enhanced the pH modifying effect of the carbonate, which outweighed the destabilizing effect of crystallinity disruption caused by wet granulation (Fig. 28.5). The loss of crystallinity appeared to have no negative effect on stability when the pH was not sufficiently acidic to promote the acid-catalyzed reaction.[82] Similarly, improved stability of a formulation containing pH modifier was also reported for DMP 754 tablets and capsules when a wet granulation process was used for product manufacture. Disodium citrate was used as the pH modifier and was dissolved in the binder solution used for wet granulation. Both

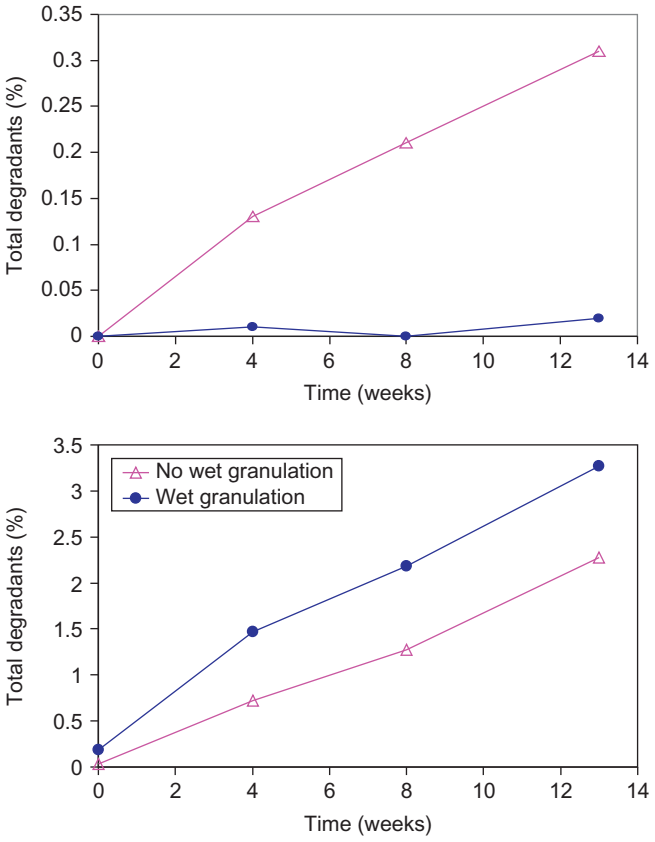


FIGURE 28.5 Formulation-dependent effect of wet granulation on stability of compound I tablets at 40°C/75% RH. Formulation containing sodium carbonate (top) and formulation without sodium carbonated (bottom). Wet granulation enhances distribution of the pH modifier, which outweighs its negative effect on crystallinity. *Source: From Badawy S, Vickery R, Shah K, Hussain M. Effect of processing and formulation variables on the stability of a salt of a weakly basic drug candidate.* Pharm Dev Technol *2004;**9**(3):239–245.*

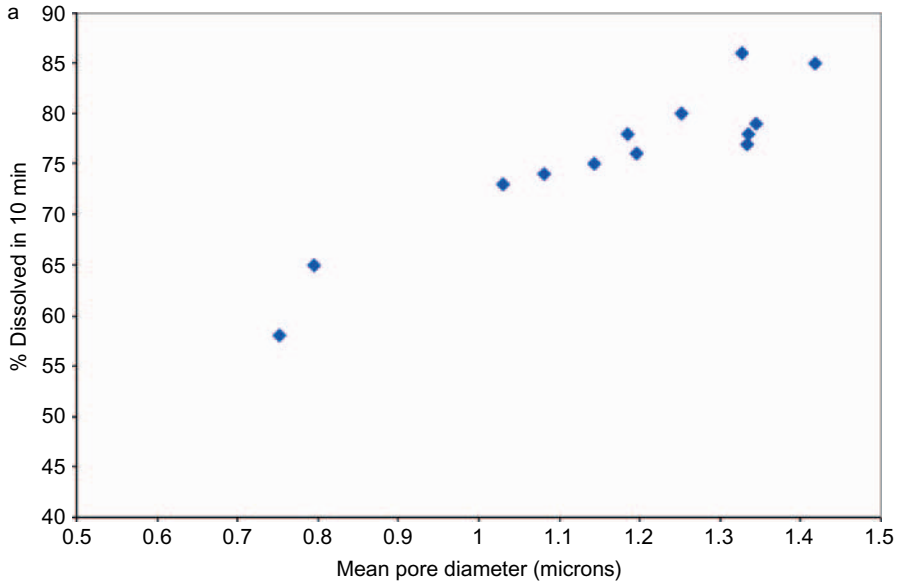


FIGURE 28.6 Effect of granule pore diameter on dissolution rate of brivanib tablets. *Source: From Badawy SI, Narang AS, LaMarche K, Subramanian G, Varia SA. Mechanistic basis for the effects of process parameters on quality attributes in high shear wet granulation.* Int J Pharm *2012;**439**(1–2):324–333.*

ester and amidine hydrolysis rates of DMP 764 were shown to decrease in the wet granulated dosage form compared with the dry granulated product. This was also attributed to the more effective pH control due to the uniform distribution of the pH modifier, which was confirmed by an x-ray imaging technique of the tablets. It is noteworthy that the improved stability of the wet granulated product in this case was observed despite the increased formulation hygroscopicity.[75]

28.5.3 Dissolution and bioavailability

Agglomeration of drug substance particles in wet granulation can potentially affect dissolution rate. Agglomerated particles have lower surface area available for dissolution and thus, dissolution rate would be reduced by granulation, unless granules disintegrate rapidly upon contact with the dissolution medium. For slowly disintegrating granules where granule erosion is the predominant mechanism for dissolution, dissolution rate is dependent on granule size. On the other hand, for rapidly disintegrating granules, granule size usually has no effect on dissolution. Disintegration of granules, however, is affected by porosity and thus, in this case, dissolution rate may be dependent on granule porosity. As porosity is lowered, granule disintegration and subsequently dissolution rate slow down. As granule pore diameter is decreased, dissolution medium ingress into the granule and subsequent granule disintegration becomes slower, resulting in the slower tablet dissolution. Thus, depending on the predominant mechanism of dissolution, granule porosity and/or size distribution may impact drug product dissolution rate. The dissolution rate of brivanib tablets was strongly dependent on granule porosity. The amount dissolved in 10 min showed a strong correlation with granule pore diameter (Fig. 28.6) and poor correlation with granule mean diameter. To the contrary, pexacerfornt tablet dissolution correlated with granule mean diameter and showed no dependence on blend density.[51] As both granule size and porosity are affected by granulation parameters and starting material and granulating liquid properties, wet granulation process can thus be modulated to optimize dosage form dissolution. Depending on the biopharmaceutical properties of the drug molecule, the effect of granulation process on dissolution rate may also impact bioavailability.

Apart from the impact of wet granulation on dissolution rate through its effect on granule size and structure, wet granulation may also affect dissolution rate of primary drug substance particles. The intimate contact of drug substance and excipients achieved in wet granulation can be leveraged to increase effectiveness of surfactants as wetting agents. For compounds with dissolution rate limited bioavailability, the resulting improvement in drug particle wetting by the dissolution medium can enhance dissolution rate and bioavailability. In a recent article, Pandey et al. showed that for a formulation exhibiting a dramatic positive food effect, the type and level of surfactant and binder were able to modulate the dissolution rates

and, in turn, the bioavailability.[3,4] The incorporation of a surfactant such as Vitamin E TPGS and poloxamer via a high-shear wet granulation process was found to significantly enhance dissolution of the tablet formulation. The study data showed that the presence of Vitamin E TPGS in dogs was essential in enhancing the bioavailability in the fasted state, and this was confirmed in human subjects where this formulation was successful in mitigating the food effect.[3]

28.5.4 Compaction and flow properties

Wet granulation can have considerable effect on compaction properties of the resulting granulation. Change in compaction properties (tablet tensile strength vs applied compression pressure) upon wet granulation is a combined effect of loss in porosity and size enlargement of the constituent particles, and the concomitant increase in the bonding and plasticity of particles upon uniform application of the binder. While the former is known to decrease compactability following wet granulation, the latter enhances the compactability through activation of the binder. In their study of a group of 25 pharmaceutical compounds and excipients, Vemavarapu et al.[29] concluded that the latter mechanism is more dominant than the former with more compounds/excipients showing enhanced compactability upon wet granulation. The increase in compactability upon wet granulation was found to be highest in materials with low initial compactability, which are more likely to gain compatibility benefit from the application of a uniform binder layer in wet granulation. To the contrary, more hygroscopic materials with high surface area resulting from high internal porosity showed decreased compatibility after wet granulation. These materials are more likely to have diminished intraparticle porosity as the result of high-shear wet granulation and thus, deterioration in compactability. Microcrystalline cellulose is a notable example of these materials. The change in pore structure of microcrystalline cellulose was shown to correlate with reduced compatibility in wet granulation, even without the increase in particle size.[83]

Despite the enhancement in compactability observed for many compounds under the granulation conditions used by Vemavarapu et al., it should be emphasized that excessive granule densification can greatly diminish compactability. Growth of overly densified granules results in reduced volume reduction behavior and compressibility (density vs compression pressure) of the granulated material.[55,84–86] Thus, many studies reported decrease in compactability with the increase in amount of granulating liquid and wet massing time and at higher impeller speed—all are factors which were shown to concomitantly decrease granule porosity.[51,52,65,87]

Improving blend-flow properties is one of the main objectives of wet granulation. The improved flow properties in wet granulation is usually the result of increased particle size and/or density of the blend. The contribution of increased size versus density was shown to vary among different formulations. The enhanced flow rate for brivanib by wet granulation was mainly attributed to increased granule density with little contribution from particle size enlargement. On the other hand, the increase in particle size (with reduction in fines) appeared to be the main mechanism for improvement in flow properties for pexacerfont formulation. The increase in both particle size and density contributed to reduced angle of repose and enhanced flow properties of razaxaban.[51]

28.6 SCALE-UP OF THE HIGH-SHEAR WET GRANULATION PROCESS

The scale-up of high-shear wet granulation process is highly relevant to the pharmaceutical and food industries. This topic has been a subject of extensive research over the years and many publications exist to date proposing different ways to scale-up this process. This section discusses the main challenges associated with scale-up and broadly categorizes the various scale-up methodologies into two main categories or strategies (Fig. 28.7).

28.6.1 Challenges in scale-up of high-shear wet granulation

The challenges associated with scale-up of high-shear wet granulation process arise from the short processing times (typically ~3–5 min), and relatively high-intensity process conditions in terms of shear, particle velocities, and particle collision frequency and energy, that are inherent to this process.[87] Change in drug substance material properties throughout the product development life cycle presents an additional challenge to scale-up. In a quality by design paradigm, there is a need for an increased understanding of the process and to establish a design space of process parameters and material properties. If such experiments are done at the large scale, it would lead to significantly higher costs. Therefore, it is important to establish the scalability of the process and be able to conduct more experimentation at the lab and/or pilot scales.

The underlying strategy during scale-up is to ensure that particles experience the same overall process

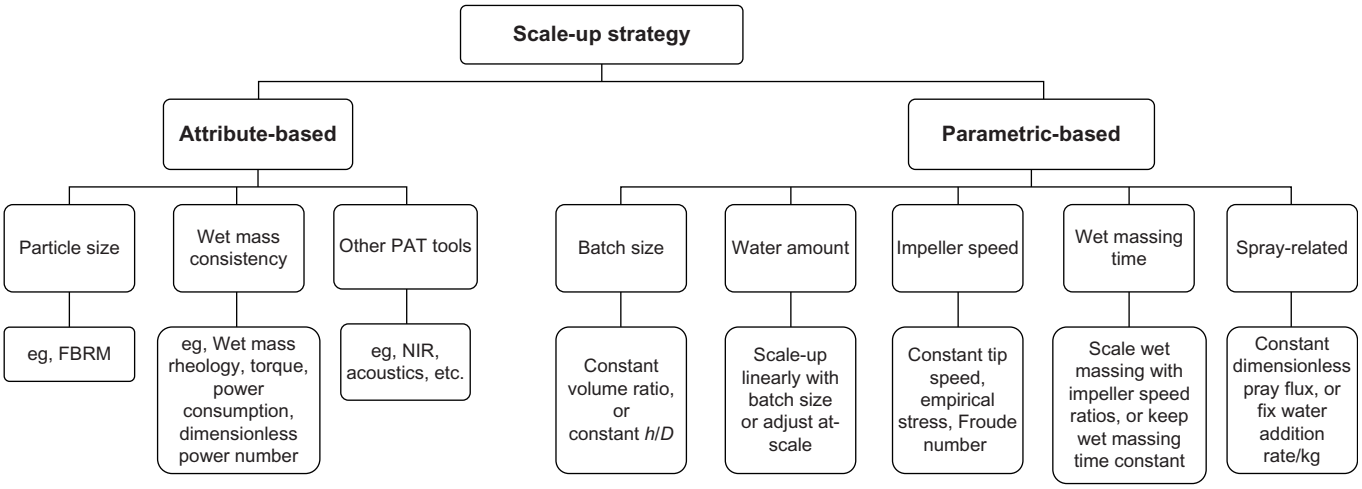


FIGURE 28.7 Scale-up strategy for high-shear wet granulation process. In parametric-based scale-up, all the parameters must be scaled-up simultaneously based on certain rules (as shown) in order to achieve success during scale-up. *Source: From Pandey P, Badawy S. A quality by design approach to scale-up of high-shear wet granulation process.* Drug dev ind pharm *2015;**42**(2):175–89.*

conditions across scales so that the same granule quality can be achieved. In order to accomplish that, there are three types of similarities required across scales: geometric, dynamic, and kinematic similarity.[18,88,89] In relation to high-shear wet granulation process, the geometric similarity refers to the granulator bowl geometry, including impeller design, and can be maintained if the ratio of linear dimensions across scales is maintained. The dynamic and kinematic similarities refer to the forces experienced by the particles inside the granulator and the particle velocities, respectively. The forces or shear experienced by the particles inside the granulator result mainly from particle-particle, particle-wall, and particle-impeller collisions. The impeller design and speed generally dictate the magnitude of shear experienced by the particles, which in turn governs the interplay between the various granulation mechanisms, such as aggregation, consolidation, and breakage (discussed in previous sections).

It is challenging to maintain geometric, dynamic, and kinematic similarities across different scales of high-shear granulators. Geometric similarity is a prerequisite to obtaining the other two similarities, and is easier to maintain if the same granulator design is used. In such cases, batch size is increased proportionally to granulator volume, which results in the same powder-bed-height-to-bowl-diameter ratio. Equipment manufacturers have recognized the importance of granulator design effects during scale-up in recent years and started offering granulators that can be set-up with different size granulators that have the same design. However, this may not be the case at the commercial scales, where routinely the granulator manufacturers are different from lab and pilot-scale manufacturers, or have a different granulator design. Sometimes the difference can be as drastic as having a top-driven versus a bottom-driven granulator, both of which will have significantly different particle flow patterns. Even though lab and pilot-scale equipment often get upgraded with new capabilities (such as monitoring torque profiles), this is generally not the case at the commercial site, where good manufacturing processes regulations are in effect and changes to equipment could be a significant undertaking, both in terms of costs and procedures. The geometrical differences across granulators result in different distribution profiles of particle velocities and forces at the different scales, which make the goal of attaining dynamic and kinematic similarities more complicated. In addition, it is also challenging to simultaneously obtain similar peak particle velocity and force across scales. As a result of these challenges, the concept of granulation "endpoint" is quite common in high-shear wet granulation. In this case, granulating liquid amount and/or wet massing time are based on an in-process measurement, which is used to determine when the granulation reaches desired properties (or endpoint) and consequently the process is stopped when this endpoint is achieved.[18]

In looking at the various scale-up rules proposed in the literature, it is observed that the scale-up methodologies can be categorized broadly into two main strategies: (1) parametric-based scale-up strategies in which target parameters identified at the small scale and shown to provide the desired final product attributes are scaled up using certain empirical calculations or; (2) attribute-based scale-up strategies in which parameters are adjusted at the different scales to achieve target in-process attributes or endpoint,

as determined by an in-process measurement.[18] A combination of the two approaches is often utilized in practice. A more detailed discussion on each of these categories is provided in the subsequent subsections.

28.6.2 Scale-up principles

28.6.2.1 Parameter-based process scale-up strategies

A parametric-based scale-up strategy is one in which the process parameters at the larger scale are determined from the small-scale parameters using certain scale-up principles for each or many of those parameters. These scale-up factors are most commonly related to the impeller speed, wet massing time, water amount, and sometimes to the granulating liquid addition (eg, spray rate, water addition time, droplet size, mode of addition). The first step in a parametric-based approach is to conduct small- or pilot-scale experiments in order to help identify CPPs that govern the granulation properties for that formulation and then scale-up those critical parameters appropriately.[61,90,91] Scale-up rules for the frequently used CPPs are listed next.

28.6.2.1.1 Water Amount

The water amount can be considered as the most critical parameter that governs the granulation process. It can be sometimes considered as a formulation factor, but given that water is introduced first and then taken out at the end (drying), it acts more as a process aid/parameter. The general strategy for scale-up of water amount is to increase it proportionally to the batch size at the large scale, in order to keep the ratio of water to intra-granular solids constant at the different scales.[92] This approach makes sense because the water amount required should be governed by the formulation components (powder properties), which do not change during scale-up. If other parameters have been scaled appropriately, there should be no need to make further adjustments in the water amount at the larger scale. However, amount of water is sometimes adjusted to compensate for other changes during granulation.[87,90] The amount of water may need to be adjusted at the large scale in some cases in order to achieve similar growth and consolidation across scales because it is usually not possible to exactly match particle velocity and force distributions at the different scales. Altering water amount at the large scale becomes an iterative approach and demands such experiments be conducted at the large scale, which can be costly. The general understanding is that a higher amount of water may be needed at larger scales to match granule properties across scales.[87]

28.6.2.1.2 Impeller Speed

Impeller speed is directly related to the shear that is applied to the granules, and also governs the flow pattern of the powder in the granulator. There are numerous publications that discuss impeller speed–related scale-up rules.[2,92–96] The most commonly used methods of scaling impeller speed use a power law correlation, as shown in Eq. (28.11).

$$\frac{rpm_2}{rpm_1} = \left(\frac{D_1}{D_2}\right)^n \tag{28.11}$$

where rpm_2 and rpm_1 are the impeller speeds in revolutions per minute, D_1 and D_2 are the impeller diameters in the two granulators, and n is the power law number. The three most commonly used values of n are 0.5, 0.8, and 1. A value of $n = 0.5$ corresponds to the case when Froude number across the scales is kept constant. In a high-shear granulator, the Froude number is essentially the ratio of angular acceleration to the gravitational acceleration. A constant Froude number maintains similar forces at the different scales, which would aid in maintaining dynamic similarity across scales. A value of $n = 1$ corresponds with keeping the tip speed constant across scales and is the more commonly used value of n, which aims to keep similar particle velocity (kinematic similarity) across scales. A constant tip speed aims to maintain constant shear rate and particle velocity at the tip of the impeller across scales.[18]

An experimental study conducted by Tardos et al.[95] suggested that neither approach matches average shear force experienced by the granules. The study used tracer pellets to measure the stress experienced by the granules across scales and suggested that a value of $n = 0.8$ results in similar shear stress across different scales. Thus, empirical matching of the stress in the granulator based on this study results in large-scale impeller speed between constant tip speed and constant Froude number approaches. This empirically derived value of n (constant empirical shear stress) was used successfully to scale-up the wet granulation process.[44,94,95] Another scale-up parameter found in the literature that is related to impeller speed is known as "relative swept volume."[34,94,97–101] It is defined as the volume swept by the impeller per second divided by the volume of the granulator. This is associated with the work input on the material that provides densification of the consolidated mass.[91]

There is disparity in literature on which impeller speed scale-up rule works best for high-shear wet granulation. One of the limitations when using Froude number, which is more fundamental in nature, is the constraints with the equipment itself. Depending on the scale difference, this rule may predict impeller

speed values that are not achievable with certain granulators. Therefore, a practical limitation with this rule often limits the use of Froude number. The most commonly used rule is where tip speed is kept constant across scales. However, there is existing literature that suggests that a constant tip speed will not ensure similar mixing patterns inside the bowl, which means that the granulation process will be different.[22,95,102] In a recent study by Tao et al., the commonly used impeller speed scale-up rules were evaluated using a microcrystalline cellulose—lactose-based low drug loading formulation with a comprehensive characterization of granule properties.[87] It was shown that the constant tip speed rule produced slightly less granulated materials at the larger scales and that a longer wet massing time at the larger scale can be used to produce granules of comparable properties across scales. Constant Froude number and constant empirical stress rules yielded granules that were more comparable across different scales, but it was pointed out that there may be practical equipment constraints limiting the use of constant Froude number approach, especially if there is a big scale-up factor between the small and large scales.

28.6.2.1.3 Wet Massing Time

There is no universal rule of how wet massing time should be scaled-up. Typical wet massing time at the lab and pilot-scale range from 15 s to 1 min, and 2−4 min at the commercial scales. A constant wet massing time across scales is sometimes used. Alternatively, wet massing time can be adjusted based on a ratio of impeller speeds from one scale to the other in order to keep the same total number of impeller rotations across scales.[92] A similar number of impeller revolutions across scales are intended to maintain constant residence time of the particles in the higher shear regions. If the impeller speed is scaled up by constant tip speed, then the constant number of impeller rotations rule would result in wet massing time that is proportional to the granulator bowl diameters. Limitations may exist for a constant number of impeller rotations rule when scaling between granulators that vary greatly in size. This approach implicitly assumes that granule growth and consolidation is predominantly limited to the regions of peak force in the granulator and thus, the same number of rotations across scales aims to maintain similar number of "passes" in these peak force regions. An opposite extreme would be to keep wet massing time constant at the different scales, assuming a narrow distribution of forces within the granulator.[18] In some cases, wet massing time is used as a means to reach a desired endpoint during scale-up, as described next under the attribute-based approach. This means that the wet massing phase is continued until a desired set of granule properties is reached. This method may ensure that some of the granule properties are matched between scales, but this may not be a desirable approach given that an extended wet massing time is generally detrimental to most of the important granule properties, such as granule porosity.[87]

28.6.2.1.4 Spray-related Parameters

There are factors to be considered during scale-up in terms of water addition, namely liquid addition time or liquid addition rate and nozzle properties (which dictates droplet size, spray area, droplet velocity, etc.). Some of these spray-related factors can be combined into a single parameter: the dimensionless spray flux described in Section 28.2.1. Given that dimensionless spray flux is not a scale-dependent property, it has been a subject of several scale-up studies as a parameter to be kept constant during scale-up.[102,103] It can serve as a good mechanistic basis for scale-up of spray-zone conditions in order to achieve equivalent liquid distribution and nucleation across scales. It has also been shown that at low dimensionless spray flux, a narrow nuclei size distribution is obtained.[22] However, in practice, pharmaceutical high-shear wet granulation usually operates in the mechanical dispersion regime in the production scale equipment. While it may be possible to operate in the droplet-controlled regime at the small scale using an appropriately selected nozzle and spray rate, scale-up of such a process is usually not feasible from a practical perspective. If the liquid addition time is kept constant across scales, then the liquid spray rate has to increase significantly when scaling up. This means that the dimensionless spray flux would increase significantly, which would shift the nucleation regime. On the other hand, if spray rate is kept constant, then it would lead to a significant increase in liquid addition time, which can affect the granule agglomeration and consolidation process, in addition to increasing the overall process time. A way to handle this would be to increase the number of spray nozzles used,[104] which may not always be a practical solution, and would require re-engineering of the granulator bowl. In addition, batch size generally increases by volume during scale-up and it would be difficult to proportionally increase the number of spray nozzles. Another challenge with the spray flux approach is the need to have full characterization of the spray properties (eg, spray area, droplet size).[105]

A practical approach to scale-up of liquid addition in the mechanical dispersion regime has been to keep liquid addition time constant, thus increasing spray rate in proportion to batch size. This usually works well if there is sufficient mixing and mechanical

forces to dispense the liquid at the large scale.[18] If this is not the case, then local overwetting stemming from poor liquid distribution can be observed at the larger scale where liquid/solid mixing is less efficient, which results in large, oversized granules and wide size distribution. Therefore, as a general practice, water is often added using a nozzle (with or without atomization) that helps to spread the water over a broader area to reduce local overwetting. This is especially important if the formulation contains highly water-soluble components.

28.6.2.2 *Attribute-based process scale-up: granulation endpoint*

An attribute-based scale-up strategy focuses on granule attributes or properties rather than the process parameters themselves. In this approach, the equipment process parameters are adjusted such that equivalent granule attributes are obtained across scales.[18] An in-process in-line or at-line measurement using a process analytical technology (PAT) tool is typically used to monitor progress of the granulation and to guide process parameters, mainly granulating liquid amount and/or wet massing time. In other words, an in-process granulation attribute is used to provide feedback control on process parameters and eventually dictate granulation endpoint.

There are various challenges when using the attribute-based endpoint approach for scale-up. Given that there are several in-process properties that can be potentially monitored, the key to using this scale-up strategy successfully is to first identify the appropriate endpoint measurement that ensures desired product quality for that particular formulation/system. This can be formulation-specific and is governed by the desired quality attributes for a particular drug product. In some cases, using one attribute for endpoint determination is not sufficient to ensure that the final product meets all the desired quality characteristics; a situation which may undermine the feasibility of an attribute-based scale-up strategy.

Endpoint measurements have been studied extensively, but the findings vary greatly on the best way to determine granulation endpoint. Some endpoint methods published in the literature include near-infrared, particle size by focused-beam reflectance microscopy (FBRM), power or torque, acoustics, wet-mass rheology, imaging, etc.[31,32,106–119] One of the most commonly used granule properties when using this type of scale-up strategy is particle size distribution. In-process measurement of granule size distribution using FBRM has been recently utilized.[120] FBRM probe measures chord length of particles that pass through a spinning laser beam and therefore tracks real-time changes in particle size and distribution in the granulation process. FBRM technology (Lasentec, C35 probe, Mettler-Toledo) gained attention over the past few years with the development of a modified probe with a mechanical scraper on the sapphire window that prevents probe fouling. Such real-time monitoring of the process provides a great opportunity for scale-up purposes and for matching granulation kinetics across scales to ensure successful scale-up, and also to enhance process and equipment understanding.

There are other existing PAT techniques to monitor the granulation kinetics that have gained more attention recently, such as capacitive sensor, strain-gauge torque sensor, near-infrared (NIR), Raman, acoustic monitoring, etc.[13,111,119,121–127] Rantanen et al. and Luukkonen et al. both demonstrated the utility of an in-line NIR spectroscopy to high-shear wet granulation process.[124,126] Rantanen et al. concluded that this technique can be valuable for scale-up, but should be used in combination with other measurements, such as power consumption and torque, to provide full characterization of the process because NIR does not measure directly the properties of the wet mass. Acoustic monitoring is another monitoring technique that is shown to provide a signature of the granulation process. It is noninvasive in nature and can be correlated with dynamic changes in granule particle size, flow, and compaction properties.[111,123] These PAT techniques provide valuable real-time information about the progression of the granulation process and can be used for process endpoint determination and control.

Another approach for endpoint determination of the wet granulation process is related to the rheological or "consistency" changes of the wet mass during granulation. There are several ways in which this is accomplished in the literature. In some cases, this was done "at-line" by taking a sample of the wet mass from the granulator and using a mixer torque rheometer to determine sample "viscosity."[128–131] In-line measurements related to wet mass consistency are also common by using an instrumented impeller to measure impeller torque, impeller power, or other derived parameters based on impeller power (eg, work done) in order to define the granulation endpoint and guide scale-up.[96,98,106,108,114,115,132–134] Similar to the other previously mentioned PAT tools, this method captures the kinetics of the granulation process, and the data have often enabled researchers to identify different phases of granulation.

A major limitation of using a power-based approach is that the motor power differs between different equipment and can even change for a given equipment over time.[34,35,112] Therefore, some normalization techniques, such as normalized impeller work, were

proposed by researchers to account for differences in scale and impeller blade design, which can be different between granulators during scale-up.[106,133] Dimensional analysis has also been applied to the scale-up of the wet granulation process. One dimensionless number derived from such analysis is the Newton power number shown in Eq. (28.12):

$$N_p = \frac{\Delta P}{\rho \omega^3 R^5} \tag{28.12}$$

where ΔP is the power consumption (corrected by the baseline power during dry mixing stage), ρ is the wet mass density, R is the impeller radius, and ω is the rotational speed. The Newton power number can be expressed as a function of other dimensionless numbers such as Froude number, pseudo-Reynolds number, and fill or geometric ratio[62,96] as shown in Eq. (28.13).

$$\log_{10}(N_p) = a.\log_{10}(\varphi Re.Fr.fillratio) + b \tag{28.13}$$

where a and b are scale-independent regression constants, φRe is the pseudo-Reynolds number (defined by Eq. (28.14)), Fr is the Froude number, and fill ratio is the granulator volume to wet mass volume ratio.

$$\varphi Re = \frac{\rho R^2 \omega}{\mu} \tag{28.14}$$

where μ is the wet mass viscosity.

The validity of Eq. (28.13) has been shown across several scales and granulators.[108,135,136] In order to use this scale-up approach, experimental data need to be generated at different scales using geometrically similar granulators and different granulation times (or endpoints), which result in granulations with different wet mass density and viscosity. Once the relationship in Eq. (28.13) is verified and values of a and b are established for a particular formulation, this dimensionless relationship can be used to predict the power endpoint at a larger scale for the desired wet mass viscosity.[18] In absence of wet mass density and viscosity data, impeller torque measurements can be used to estimate the pseudo-Reynolds number. The Froude number during scale-up can either be kept constant or estimated if the impeller speed for the larger scale is determined by another impeller speed–related scale-up rule (eg, constant tip speed) as described previously.

In practice, often a combination of the two scale-up strategies (attribute-based and parametric-based) is utilized, with certain process parameters being scaled-up using the discussed scale-up principles (eg, impeller speed based rules), but also controlling some of the process parameters (eg, wet massing time) such that the desired endpoint or equivalent granule properties (eg, particle size and porosity) are achieved.

28.7 MODELING AND SIMULATION IN HIGH-SHEAR WET GRANULATION

There has been increased interest in modeling and simulation of wet granulation in recent years. Wet granulation models can potentially be utilized to increase understanding of the effect of process parameters on resulting product attributes and can aid in process scale-up. A fully predictive model can also be used for process control. At this time, high fidelity predictive models of pharmaceutical wet granulation are not yet available and, therefore, existing models are predominantly applied to enhance mechanistic understanding during process design and scale-up, rather than process control. The various modeling approaches utilized for high-shear wet granulation process include regime-map based models, population balance modeling (PBM), discrete element modeling (DEM), PBM coupled with DEM, PBM coupled with computational fluid dynamics, PBM with compartmental model and DEM, and PBM coupled with volume of fluid models.[65,137–145] Kayrak-Talay et al. provided a recent review of regime-map based models for *a priori* design and scaling.[146] This section discusses the more commonly used PBM and DEM-based approaches, and the more recently used hybrid approach between PBM and DEM.

28.7.1 Population balance modeling

PBMs are number balance models around each granule size (volume) fraction where the change in the number of granules in a given size is equal to the number of granules formed minus the number of granules leaving the size fraction. Particles in any size fraction are assumed to interact with each other and with the environment, which results in the change in the number frequency of particles in the size fraction over time. For a wet granulation system with only aggregation (coalescence) rate process, the PBM to describe the system can be written as[147]:

$$\frac{\delta n(u,t)}{\delta t} = B_{coal}(u,t) - D_{coal}(u,t) \tag{28.15}$$

where $n(u,t)$ is the number frequency of granules of size u at time t. $B_{coal}(u,t)$ and $D_{coal}(u,t)$ represent the rate of formation and disappearance of granules of size u due to the binary coalescence of granules of size v with u-v and u with v, respectively. These are given by

$$B_{coal}(u,t) = \frac{1}{2N(t)} \int_{v=0}^{u} \beta^*(v, u-v) n(v,t) n(u-v,t) dv \tag{28.16}$$

$$D_{coal}(u,t) = \frac{1}{N(t)} \int_{v=0}^{\infty} \beta^*(u,v) n(u,t) n(v,t) dv \qquad (28.17)$$

where $\beta^*(u,v)$ is referred to as the coalescence kernel, which is the normalized frequency of collisions between granules of sizes u and v that result in coalescence, and $N(t)$ is the total number of granules per unit volume in the granulator. Similarly, $\beta^*(v, u\text{-}v)$ is the coalescence kernel of granules v and u-v. Early work by Annapragada and Neilly[148] demonstrated that PBM based on aggregation term alone does not adequately describe experimental data for wet granulation and proposed the addition of breakage term to the PBM equation, such as the one proposed by Ramachandran et al.[145] In addition to breakage, other terms can be also added to Eq. (28.15) to describe all relevant rate processes in the granulator that result in creation or death of granules, as shown in Eq. (28.18).

$$\frac{\delta n(u,t)}{\delta t} = B_{nuc}(u,t) + B_{coal}(u,t) - D_{coal}(u,t) + B_{break}(u,t) - D_{break}(u,t) \qquad (28.18)$$

where B_{nuc} is the rate of nucleation and $B_{break}(u, t)$ and $D_{break}(u, t)$ represent the rate of formation and disappearance of granules of size u due to breakage.

The coalescence kernel $\beta^*(u,v)$ is at the core of the PBM model. The value of the PBM depends on the success of the kernel to accurately define coalescence in the system. There are numerous literature reports describing different expressions for $\beta^*(u,v)$.[142,144,149–151] Earlier reports used empirical kernels where the kernel form was assumed and its value was estimated by fitting data from granulation studies to the model.[152–154] The validity of the kernel was concluded if the model provided good fit to the experimental data. Such an empirical approach has significant limitations due the kernel's lack of physical relevance and because predictions can be made only for the same scale and within a certain process space spanned by the experimental design. Recently, development of physical-model-based kernels is becoming more common where microscale (particle level) models of granulation are used to define the kernels, thus increasing the scope and applicability of PBM models. Kernels based on models that account for these factors are expected to be more successful in describing the granulation process because granule and liquid properties as well as granule collision velocity dictate probability of coalescence. Earlier attempts to establish physical-based kernels were limited to invoking physical models to only select kernel form, but were not predictive in nature as rate constants were still obtained by data fitting. For example, Liu and Litster[151] proposed a coalescence kernel the value of which depends on the type of coalescence (type I or II). Adetayo and Ennis[149] defined a critical granule size based on Stokes viscous number. All granule collisions where the effective mean granule size of the colliding granules is less than the critical size were considered successful and modeled by a size-independent kernel. On the other hand, collisions with mean granule size greater than the critical size were considered unsuccessful with zero rate of growth. Ideally, models with physical-based kernels should be predictive with no fitted parameters where the kernel is completely defined *a priori* based on physical models of coalescence. Physical parameters in the model are preferably determined by independent calibration experiments, rather than fitting of granulation study data to the PBM. More recent PBM models have moved in this direction,[65,138,141,144,155] but kernels still include some fitting parameters that make the PBM not entirely mechanistic in nature. Accounting for the particle collision velocities in the coalescence kernel still remains as one of the key challenges due to the wide and scale-dependent distribution of particle velocities in the granulator. One approach to address this challenge would be to model the granulator as several segregated regions with different collision velocities in each region and size-dependent rates of transfer of granules between each zone.[147] Once model development is completed, model performance should be verified using new studies independent of the data set used to establish the model.

The coalescence kernel is often split into two components[150,156]:

$$\beta^*(u,v) = C_{u,v}\beta(u,v) \qquad (28.19)$$

where the first term $C_{u,v}$ is related to collision frequency of particles u and v and while the second term, $\beta^*(u,v)$, is a function of collision efficiency (ie, the probability that a collision of the two particles results in successful coalescence). $C_{u,v}$ accounts for process parameters, such as impeller speed and some granule properties such as granule size. $\beta^*(u,v)$ has traditionally been dependent on granule size alone in the one-dimensional PBM models where the probability of successful coalescence is assumed to depend only on particle size of colliding granules. However, physical models of coalescence clearly show that probability of coalescence is dependent on other granule properties, such as liquid content and porosity. Therefore, a fully predictive PBM model needs to include multidimensional granule distributions of size, porosity, and liquid content as independent

variables that evolve with time.[147] A three-dimensional PBM is expressed as:

$$\frac{\delta n(m,\varepsilon,l,t)}{\delta t} = B_{\text{coal}}(m,\varepsilon,l,t) - D_{\text{coal}}(m,\varepsilon,l,t) + C(m,\varepsilon,l,t) + L(m,\varepsilon,l,t) \tag{28.20}$$

where m is the mass of granule solid phase, l is granule liquid to solid ratio, and ε is granule porosity. In addition to the coalescence terms, the equation also includes terms for evolution of porosity, C, and liquid content, L, due to consolidation and liquid addition, respectively. The coalescence terms are expressed as

$$B_{\text{coal}}(m,\varepsilon,l,t) = \frac{1}{2N(t)}\int_{m_1=0}^{m}\int_{\varepsilon_1=0}^{1}\int_{l_1=0}^{\infty}\beta^*(m_1,m_2,\varepsilon_1,\varepsilon_2,l_1,l_2)\, n(m_1,\varepsilon_1,l_1,t)n(m_2,\varepsilon_2,l_2,t)dm_1d\varepsilon_1dl_1 \tag{28.21}$$

$$D_{\text{coal}}(m,\varepsilon,l,t) = \frac{n(m,\varepsilon,l,t)}{N(t)}\int_{m_1=0}^{\infty}\int_{\varepsilon_1=0}^{1}\int_{l_1=0}^{\infty}\beta^*(m_1,m,\varepsilon_1,\varepsilon,l_1,l)\, n(m_1,\varepsilon_1,l_1,t)dm_1d\varepsilon_1dl_1 \tag{28.22}$$

where granule with property m,ε,l is created by coalescence of granules m_1,ε_1,l_1 and m_2,ε_2,l_2 and disappears by coalescence with granule m_1,ε_1,l_1. The resulting is a complex, three-dimensional integro-differential equation which is computationally expensive to solve. While conventional numerical methods have been used to solve the multidimensional PBM equations, more efficient computational methods such as Monte Carlo techniques have been recently utilized to reduce computation times.[142,157,158]

One of the utilities of PBM is the ability to conduct a sensitivity analysis of some of the main formulation and process parameters (eg, viscosity, impeller speed, contact angle, and spray rate) using minimal experiments.[159] Such a sensitivity analysis can be useful in establishing a process and formulation design space of a drug product. PBM predictions can also be used to reduce experimentation once a good model is established using some initial experimentation that is required for determining the kernel parameters for the system being modeled.[65,140]

28.7.2 Discrete element modeling

DEM is a numerical simulation technique which tracks the position of every particle within the defined geometry of the granulator and estimates the velocity and force exerted on every particle.[18] In contrast to PBM models, which are meso- and macro-scale models, DEM models bridge the gap between the micro- and meso-scale and have some advantages over PBM simulations. DEM provides detailed information about the particle dynamics, such as time evolution of individual particle velocities and shear experienced by the particles.[140,143,160–162] Two distinct DEM methods have been used, which are the hard sphere and soft sphere methods. While the hard sphere method is computationally less demanding, it is generally less applicable to the pharmaceutical granule due to the porosity and liquid content associated with granule microstructure. The soft sphere method is more capable of simulating particles in pharmaceutical wet granulation, but is computationally more expensive and capable of handling fewer particles than the hard sphere method.[142] As a result, DEM simulation of a pharmaceutical wet granulation system is computationally demanding and not currently capable of simulating the large number of particles present in the high-shear granulator. DEM simulation is usually carried out using larger particle size (in the few mm range[143]) than what is typically used in the pharmaceutical granulation because DEM can only handle few hundred thousand to a million particles. Calibration of microscale particle properties also presents another challenge to DEM. As result of these complexities, some literature reports used DEM to study a limited aspect of wet granulation, rather than applying it for complete prediction of the wet granulation process. DEM proved to be a useful tool in studying flow patterns and force distributions in the granulator across scales.[143,150,163,164] Such applications do not provide a predictive model of wet granulation, but are still valuable in process design and scale-up. In one such example, DEM was used to understand differences in flow patterns and particle velocity distribution between different granulators and scales, which are useful from a scale-up perspective.[143]

28.7.3 Combined PBM/DEM approach

A combined approach using both PBM and DEM has been proposed to model high-shear wet granulation process because PBM and DEM are different scale models. Multiscale DEM/PBM molding is emerging as a potentially powerful approach for high-shear wet granulation.[150,165–167] Microscale DEM is used to model particle flow in the granulator and characterize particle collisions. In addition, DEM is also used to establish collision efficiency at the microscale.[143] The two feed into the PBM coalescence kernel for meso-scale simulation of wet granulation. Gantt et al. provided an

example of the combined approach [140,150] in which they integrated DEM and the physical model of coalescence in parallel for simultaneous microscale particle analysis. Relative velocities of colliding particles were calculated by DEM simulation. These velocities, together with granule porosity and liquid content, were used to determine whether coalescence will take place based on the physical coalescence model. The data obtained from the DEM/coalescence model were used to estimate the coalescence efficiency term of the PBM kernel. Simple regression was used to map particle collision and coalescence information from the DEM/coalescence model to a four-dimensional coalescence efficiency term, which is dependent on particle diameter, pore saturation, porosity, and collision velocity. DEM was also used to establish frequency and size dependence of granule collisions in the granulator and estimate the collision frequency component, $C_{u,v}$, of the PBM kernel.

28.8 SUMMARY

High-shear wet granulation is a complex unit operation that involves multiple simultaneous rate processes. In addition, the outcome of wet granulation is strongly impacted by input material and granulating liquid properties. In order to design a robust drug product, it is important to have a good understanding of the CMAs and CPPs and how they are related to the drug product's CQAs. Various material properties, such as particle size, surface area, contact angle, and solubility have profound effects on the wet granulation process together with granulating liquid surface tension and viscosity. In addition, there are usually significant interactions between formulation components and process variables dictating the granulation outcome and thus, successful design of a robust wet granulation process should account for these interactions. Key process parameters known to affect wet granulation were discussed in this chapter and their role in the design of a robust process was articulated.

Scale-up of the wet granulation process presents another challenge. A section in this chapter discussed the challenges associated with scale-up of the high-shear wet granulation process, the various scale-up principles that are commonly used, and their broad classification into attribute-based and process parameter-based approaches. In addition, modeling tools that can be used to enhance understanding of the process, and potentially allow for in silico design space development in the future were also discussed.

References

1. Nguyen TH, Morton DAV, Hapgood KP. Application of the unified compaction curve to link wet granulation and tablet compaction behaviour. *Powder Technol* 2013;**240**(0):103–15.
2. Rahmanian N, Ng B, Hassanpour A, Ghadiri M, Ding Y, Jia X, et al. Scale-up of high shear mixer granulators. *KONA* 2008; **26**:190–204.
3. Pandey P, Hamey R, Bindra DS, Huang Z, Mathias N, Eley T, et al. From bench to humans: formulation development of a poorly water soluble drug to mitigate food effect. *AAPS PharmSciTech* 2014;**15**(2):407–16.
4. Pandey P, Sinko PD, Bindra DS, Hamey R, Gour S, Vema-Varapu C. Processing challenges with solid dosage formulations containing vitamin E TPGS. *Pharm Dev Technol* 2013; **18**(1):296–304.
5. Herting MG, Kleinebudde P. Roll compaction/dry granulation: effect of raw material particle size on granule and tablet properties. *Int J Pharm* 2007;**338**(1–2):110–18.
6. Thoorens G, Krier F, Leclercq B, Carlin B, Evrard B. Microcrystalline cellulose, a direct compression binder in a quality by design environment—a review. *Int J Pharm* 2014; **473**(1–2):64–72.
7. He X, Secreast PJ, Amidon GE. Mechanistic study of the effect of roller compaction and lubricant on tablet mechanical strength. *J Pharm Sci* 2007;**96**(5):1342–55.
8. Pandey P, Bindra DS, Gour S, Trinh J, Buckley D, Badawy S. Excipient-process interactions and their impact on tablet compaction and film coating. *J Pharm Sci* 2014;**103**(11):3666–74.
9. Badawy S, Hussain M, Zhao F, Ye Q, Huang Y, Palaniswamy V. Degradation pathways of a corticotropin-releasing factor antagonist in solution and solid states. *J Pharm Sci* 2009;**98**(8):2636–47.
10. Badawy S, Vickery R, Shah K, Hussain M. Effect of processing and formulation variables on the stability of a salt of a weakly basic drug candidate. *Pharm Dev Technol* 2004;**9**(3):239–45.
11. Gift AD, Luner PE, Luedeman L, Taylor LS. Manipulating hydrate formation during high shear wet granulation using polymeric excipients. *J Pharm Sci* 2009;**98**(12):4670–83.
12. Wikström H, Carroll WJ, Taylor LS. Manipulating theophylline monohydrate formation during high-shear wet granulation through improved understanding of the role of pharmaceutical excipients. *Pharm Res* 2008;**25**(4):923–35.
13. Wikström H, Marsac PJ, Taylor LS. In-line monitoring of hydrate formation during wet granulation using Raman spectroscopy. *J Pharm Sci* 2005;**94**(1):209–19.
14. Polizzi MA, Singhal D, Colvin J. Mechanoradical-induced degradation in a pharmaceutical blend during high-shear processing. *Pharm Dev Technol* 2008;**13**(6):457–62.
15. Knight PC, Seville JPK, Wellm AB, Instone T. Prediction of impeller torque in high shear powder mixers. *Chem Eng Sci* 2001;**56**(15):4457–71.
16. Ax K, Feise H, Sochon R, Hounslow M, Salman A. Influence of liquid binder dispersion on agglomeration in an intensive mixer. *Powder Technol* 2008;**179**(3):190–4.
17. D'alonzo GD, O'connor RE, Schwartz JB. Effect of binder concentration and method of addition on granule growth in a high intensity mixer. *Drug Dev Ind Pharm* 1990;**16**(12):1931–44.
18. Pandey P, Badawy S. A quality by design approach to scale-up of high-shear wet granulation process. *Drug Dev Ind Pharm* 2015;**42**(2):175–89.
19. Rowe RC. Correlation between predicted binder spreading coefficients and measured granule and tablet properties in the granulation of paracetamol. *Int J Pharm* 1990;**58**(3):209–13.

20. Rowe RC. Binder-substrate interactions in granulation: a theoretical approach based on surface free energy and polarity. *Int J Pharm* 1989;**52**(2):149–54.
21. Iveson SM, Litster JD, Hapgood K, Ennis BJ. Nucleation, growth and breakage phenomena in agitated wet granulation processes: a review. *Powder Technol* 2001;**117**(1–2):3–39.
22. Litster JD, Hapgood KP, Michaels JN, Sims A, Roberts M, Kameneni SK, et al. Liquid distribution in wet granulation: dimensionless spray flux. *Powder Technol* 2001;**114**(1–3):32–9.
23. Hapgood KP, Amelia R, Zaman MB, Merrett BK, Leslie P. Improving liquid distribution by reducing dimensionless spray flux in wet granulation—a pharmaceutical manufacturing case study. *Chem Eng J* 2010;**164**(2–3):340–9.
24. Hapgood KP, Litster JD, Biggs SR, Howes T. Drop penetration into porous powder beds. *J Colloid Interface Sci* 2002;**253**(2):353–66.
25. Hapgood KP, Litster JD, Smith R. Nucleation regime map for liquid bound granules. *AIChE J* 2003;**49**(2):350–61.
26. Hapgood KP, Litster JD, White ET, Mort PR, Jones DG. Dimensionless spray flux in wet granulation: Monte-Carlo simulations and experimental validation. *Powder Technol* 2004;**141**(1–2):20–30.
27. Hapgood KP, Tan MXL, Chow DWY. A method to predict nuclei size distributions for use in models of wet granulation. *Adv Powder Technol* 2009;**20**(4):293–7.
28. Denesuk M, Smith GL, Zelinski BJJ, Kreidl NJ, Uhlmann DR. Capillary penetration of liquid droplets into porous materials. *J Colloid Interface Sci* 1993;**158**(1):114–20.
29. Vemavarapu C, Surapaneni M, Hussain M, Badawy S. Role of drug substance material properties in the processibility and performance of a wet granulated product. *Int J Pharm* 2009;**374**(1–2):96–105.
30. Scott AC, Hounslow MJ, Instone T. Direct evidence of heterogeneity during high-shear granulation. *Powder Technol* 2000;**113**(1–2):205–13.
31. Kristensen HG, Holm P, Schaefer T. Mechanical properties of moist agglomerates in relation to granulation mechanisms. Part II. Effects of particle size distribution. *Powder Technol* 1985;**44**(3):239–47.
32. Kristensen HG, Holm P, Schaefer T. Mechanical properties of moist agglomerates in relation to granulation mechanisms. Part I. Deformability of moist, densified agglomerates. *Powder Technol* 1985;**44**(3):227–37.
33. Iveson SM, Litster JD, Ennis BJ. Fundamental studies of granule consolidation Part 1: effects of binder content and binder viscosity. *Powder Technol* 1996;**88**(1):15–20.
34. Kristensen HG, Schaefer T. Granulation: a review on pharmaceutical wet-granulation. *Drug Dev Ind Pharm* 1987;**13**(4–5):803–72.
35. Ritala M, Holm P, Schaefer T, Kristensen HG. Influence of liquid bonding strength on power consumption during granulation in a high shear mixer. *Drug Dev Ind Pharm* 1988;**14**(8):1041–60.
36. Ritala M, Jungersen O, Holm P, Schæfer T, Kristensen HG. A comparison between binders in the wet phase of granulation in a high shear mixer. *Drug Dev Ind Pharm* 1986;**12**(11–13):1685–700.
37. Iveson SM, Litster JD. Growth regime map for liquid-bound granules. *AIChE J* 1998;**44**(7):1510–18.
38. Iveson SM, Wauters PAL, Forrest S, Litster JD, Meesters GMH, Scarlett B. Growth regime map for liquid-bound granules: further development and experimental validation. *Powder Technol* 2001;**117**(1–2):83–97.
39. Liu LX, Litster JD, Iveson SM, Ennis BJ. Coalescence of deformable granules in wet granulation processes. *AIChE J* 2000;**46**(3):529–39.
40. Ennis BJ, Tardos G, Pfeffer R. A microlevel-based characterization of granulation phenomena. *Powder Technol* 1991;**65**(1–3):257–72.
41. Ouchiyama N, Tanaka T. The probability of coalescence in granulation kinetics. *Ind Eng Chem Process Des Dev* 1975;**14**(3):286–9.
42. Ramaker JS, Jelgersma MA, Vonk P, Kossen NWF. Scale-down of a high-shear pelletisation process: flow profile and growth kinetics. *Int J Pharm* 1998;**166**(1):89–97.
43. Vonk P, Guillaume CPF, Ramaker JS, Vromans H, Kossen NWF. Growth mechanisms of high-shear pelletisation. *Int J Pharm* 1997;**157**(1):93–102.
44. Tardos GI, Khan MI, Mort PR. Critical parameters and limiting conditions in binder granulation of fine powders. *Powder Technol* 1997;**94**(3):245–58.
45. Pietsch W, Hoffman E, Rumpf H. Tensile strength of moist agglomerates. *Ind Eng Chem Product Res Dev* 1969;**8**(1):58–62.
46. Badawy SI, Hussain MA. Effect of starting material particle size on its agglomeration behavior in high shear wet granulation. *AAPS PharmSciTech* 2004;**5**(3):e38.
47. Badawy SI, Lee TJ, Menning MM. Effect of drug substance particle size on the characteristics of granulation manufactured in a high-shear mixer. *AAPS PharmSciTech* 2000;**1**(4):E33.
48. Jaiyeoba KT, Spring MS. The granulation of ternary mixtures: the effect of the solubility of the excipients. *J Pharm Pharmacol* 1980;**32**(1):1–5.
49. Rahmanian N, Naji A, Ghadiri M. Effects of process parameters on granules properties produced in a high shear granulator. *Chem Eng Res Design* 2011;**89**(5):512–18.
50. Millili GP, Schwartz JB. The strength of microcrystalline cellulose pellets: the effect of granulating with water/ethanol mixtures. *Drug Dev Ind Pharm* 1990;**16**(8):1411–26.
51. Badawy SI, Narang AS, LaMarche K, Subramanian G, Varia SA. Mechanistic basis for the effects of process parameters on quality attributes in high shear wet granulation. *Int J Pharm* 2012;**439**(1–2):324–33.
52. Badawy SI, Menning MM, Gorko MA, Gilbert DL. Effect of process parameters on compressibility of granulation manufactured in a high-shear mixer. *Int J Pharm* 2000;**198**(1):51–61.
53. Knight PC. An investigation of the kinetics of granulation using a high shear mixer. *Powder Technol* 1993;**77**(2):159–69.
54. Rahmanian N, Ghadiri M, Jia X, Stepanek F. Characterisation of granule structure and strength made in a high shear granulator. *Powder Technol* 2009;**192**(2):184–94.
55. Johansson B, Alderborn G. The effect of shape and porosity on the compression behaviour and tablet forming ability of granular materials formed from microcrystalline cellulose. *Eur J Pharm Biopharm* 2001;**52**(3):347–57.
56. Oulahna D, Cordier F, Galet L, Dodds JA. Wet granulation: the effect of shear on granule properties. *Powder Technol* 2003;**130**(1–3):238–46.
57. Marston JO, Sprittles JE, Zhu Y, Li EQ, Vakarelski IU, Thoroddsen ST. Drop spreading and penetration into pre-wetted powders. *Powder Technol* 2013;**239**(0):128–36.
58. Oka S, Kašpar O, Tokárová V, Sowrirajan K, Wu H, Khan M, et al. A quantitative study of the effect of process parameters on key granule characteristics in a high shear wet granulation process involving a two component pharmaceutical blend. *Adv Powder Technol* 2015;**26**(1):315–22.
59. Badawy SI, Shah KR, Surapaneni MS, Szemraj MM, Hussain M. Effect of spray-dried mannitol on the performance of microcrystalline cellulose-based wet granulated tablet formulation. *Pharm Dev Technol* 2010;**15**(4):339–45.
60. Ohno I, Hasegawa S, Yada S, Kusai A, Moribe K, Yamamoto K. Importance of evaluating the consolidation of granules manufactured by high shear mixer. *Int J Pharm* 2007;**338**(1–2):79–86.

61. Iskandarani B, Shiromani PK, Clair JH. Scale-up feasibility in high-shear mixers: determination through statistical procedures. *Drug Dev Ind Pharm* 2001;**27**(7):651–7.
62. am Ende MT, Bharadwaj R, García-Muñoz S, Ketterhagen W, Prpich A, Doshi P. In: Am Ende David J, editor. *Process modeling techniques and applications for solid oral drug products Chemical Engineering in the Pharmaceutical Industry*. Hoboken, New Jersey: John Wiley & Sons, Inc; 2011. p. 633–62.
63. Hapgood KP, Litster JD. Wet granulation processes. In: Am Ende David J, editor. *Chemical Engineering in the Pharmaceutical Industry*. Hoboken, New Jersey: John Wiley & Sons, Inc; 2011. p. 757–80.
64. Shi L, Feng Y, Sun CC. Massing in high shear wet granulation can simultaneously improve powder flow and deteriorate powder compaction: a double-edged sword. *Eur J Pharm Sci* 2011;**43**(1–2):50–6.
65. Pandey P, Tao J, Chaudhury A, Ramachandran R, Gao JZ, Bindra DS. A combined experimental and modeling approach to study the effects of high-shear wet granulation process parameters on granule characteristics. *Pharm Dev Technol* 2013;**18**(1):210–24.
66. Cavinato M, Andreato E, Bresciani M, Pignatone I, Bellazzi G, Franceschinis E, et al. Combining formulation and process aspects for optimizing the high-shear wet granulation of common drugs. *Int J Pharm* 2011;**416**(1):229–41.
67. Pan D, Crull G, Yin S, Grosso J. Low level drug product API form analysis - Avalide tablet NIR quantitative method development and robustness challenges. *J Pharm Biomed Anal* 2014;**89**:268–75.
68. Davis TD, Morris KR, Huang H, Peck GE, Stowell JG, Eisenhauer BJ, et al. In situ monitoring of wet granulation using online x-ray powder diffraction. *Pharm Res* 2003;**20**(11):1851–7.
69. Aguiar AJ, Krc J, Kinkel AW, Samyn JC. Effect of polymorphism on the absorption of chloramphenicol from chloramphenicol palmitate. *J Pharm Sci* 1967;**56**(7):847–53.
70. Kobayashi Y, Ito S, Itai S, Yamamoto K. Physicochemical properties and bioavailability of carbamazepine polymorphs and dihydrate. *Int J Pharm* 2000;**193**(2):137–46.
71. Morris KR, Griesser UJ, Eckhardt CJ, Stowell JG. Theoretical approaches to physical transformations of active pharmaceutical ingredients during manufacturing processes. *Adv Drug Deliv Rev* 2001;**48**(1):91–114.
72. Debnath S, Suryanarayanan R. Influence of processing-induced phase transformations on the dissolution of theophylline tablets. *AAPS PharmSciTech* 2004;**5**(1).
73. Otsuka M, Kanai Y, Hattori Y. Real-time monitoring of changes of adsorbed and crystalline water contents in tablet formulation powder containing theophylline anhydrate at various temperatures during agitated granulation by near-infrared spectroscopy. *J Pharm Sci* 2014;**103**(9):2924–36.
74. Shah KR, Hussain MA, Hubert M, Farag Badawy SI. Form conversion of anhydrous lactose during wet granulation and its effect on compactibility. *Int J Pharm* 2008;**357**(1–2):228–34.
75. Badawy SIF, Williams RC, Gilbert DL. Chemical stability of an ester prodrug of a glycoprotein IIb/IIIa receptor antagonist in solid dosage forms. *J Pharm Sci* 1999;**88**(4):428–33.
76. Blanco M, Alcalá M, González JM, Torras E. Near infrared spectroscopy in the study of polymorphic transformations. *Anal Chim Acta* 2006;**567**(2):262–8.
77. Airaksinen S, Karjalainen M, Kivikero N, Westermarck S, Shevchenko A, Rantanen J, et al. Excipient selection can significantly affect solid-state phase transformation in formulation during wet granulation. *AAPS PharmSciTech [electronic resource]* 2005;**6**(2):E311–22.
78. Tantry JS, Tank J, Suryanarayanan R. Processing-induced phase transitions of theophylline - implications on the dissolution of theophylline tablets. *J Pharm Sci* 2007;**96**(5):1434–44.
79. Badawy SI, Lin J, Gokhale M, Desai S, Nesarikar VV, LaMarche KR, et al. Quality by design development of brivanib alaninate tablets: degradant and moisture control strategy. *Int J Pharm* 2014;**469**(1):111–20.
80. Waterman KC, Adami RC, Alsante KM, Antipas AS, Arenson DR, Carrier R, et al. Hydrolysis in pharmaceutical formulations. *Pharm Dev Technol* 2002;**7**(2):113–46.
81. Narang AS, Desai D, Badawy S. Impact of excipient interactions on solid dosage form stability. *Pharm Res* 2012;**29**(10):2660–83.
82. Badawy SI, Hussain MA. Microenvironmental pH modulation in solid dosage forms. *J Pharm Sci* 2007;**96**(5):948–59.
83. Badawy SI, Gray DB, Hussain MA. A study on the effect of wet granulation on microcrystalline cellulose particle structure and performance. *Pharm Res* 2006;**23**(3):634–40.
84. Wikberg M, Alderborn G. Compression characteristics of granulated materials V. Mechanical properties of individual granules, assessed by diametral compression, in granulations with different volume reduction behaviour. *STP Pharm Sci* 1992;**2**(4):313–19.
85. Wikberg M, Alderborn G. Compression characteristics of granulated materials: VI. Pore size distributions, assessed by mercury penetration, of compacts of two lactose granulations with different fragmentation propensities. *Int J Pharm* 1992;**84**(2):191–5.
86. Wikberg M, Alderborn G. Compression characteristics of granulated materials. IV. The effect of granule porosity on the fragmentation propensity and the compatibility of some granulations. *Int J Pharm* 1991;**69**(3):239–53.
87. Tao J, Pandey P, Bindra DS, Gao JZ, Narang AS. Evaluating scale-up rules of a high-shear wet granulation process. *J Pharm Sci* 2015;**104**(7):2323–33.
88. Pandey P, Turton R, Joshi N, Hammerman E, Ergun J. Scale-up of a pan-coating process. *AAPS PharmSciTech* 2006;**7**(4):E125–32.
89. Agrawal A, Pandey P. Scale up of pan coating process using Quality by Design principles. *J Pharm Sci* 2015;**104**(11):3589–611.
90. He X, Lunday KA, Li LC, Sacchetti MJ. Formulation development and process scale up of a high shear wet granulation formulation containing a poorly wettable drug. *J Pharm Sci* 2008;**97**(12):5274–89.
91. Wehrlé P, Nobelis P, Cuiné A, Stamm A. Scaling-up of wet granulation a statistical methodology. *Drug Dev Ind Pharm* 1993;**19**(16):1983–97.
92. Rekhi G, Caricofe R, Parikh D, Augsburger L. A new approach to scale-up of a high-shear granulation process. *Pharm Technol* 1996;**20**(10):1–10.
93. Hassanpour A, Antony SJ, Ghadiri M. Modeling of agglomerate behavior under shear deformation: effect of velocity field of a high shear mixer granulator on the structure of agglomerates. *Adv Powder Technol* 2007;**18**(6):803–12.
94. Horsthuis GJB, van Laarhoven JAH, van Rooij RCBM, Vromans H. Studies on upscaling parameters of the Gral high shear granulation process. *Int J Pharm* 1993;**92**(1–3):143–50.
95. Tardos GI, Hapgood KP, Ipadeola OO, Michaels JN. Stress measurements in high-shear granulators using calibrated "test" particles: application to scale-up. *Powder Technol* 2004;**140**(3):217–27.
96. Faure A, York P, Rowe RC. Process control and scale-up of pharmaceutical wet granulation processes: a review. *Eur J Pharm Biopharm* 2001;**52**(3):269–77.
97. Holm P. Effect of impeller and chopper design on granulation in a high speed mixer. *Drug Dev Ind Pharm* 1987;**13**(9–11):1675–701.

98. Holm P, Schaefer T, Kristensen H. Granulation in high-speed mixers Part V. Power consumption and temperature changes during granulation. *Powder Technol* 1985;**43**(3):213–23.
99. Holm P, Schaefer T, Kristensen HG. Granulation in high-speed mixers Part VI. Effects of process conditions on power consumption and granule growth. *Powder Technol* 1985;**43**(3):225–33.
100. Schaefer T, Bak H, Jaegerskou A, Kristensen A, Svensson J, Holm P, et al. Granulation in different types of high speed mixers. I: effects of process variables and up-scaling. *Pharm Ind* 1986;**48**(9):1083–9.
101. Schaefer T, Bak H, Jaegerskou A, Kristensen A, Svensson J, Holm P, et al. Granulation in different types of high speed mixers. II: comparisons between mixers. *Pharm Ind* 1987;**49**(3):297–304.
102. Plank R, Diehl B, Grinstead H, Zega J. Quantifying liquid coverage and powder flux in high-shear granulators. *Powder Technol* 2003;**134**(3):223–34.
103. Litster J, Ennis B, Lian L. *The science and engineering of granulation processes*. Springer; 2004ed.
104. Rambali B, Baert L, Massart DL. Scaling up of the fluidized bed granulation process. *Int J Pharm* 2003;**252**(1–2):197–206.
105. Mort PR. Scale-up and control of binder agglomeration processes—flow and stress fields. *Powder Technol* 2009;**189**(2):313–17.
106. Campbell GA, Clancy DJ, Zhang JX, Gupta MK, Oh CK. Closing the gap in series scale up of high shear wet granulation process using impeller power and blade design. *Powder Technol* 2011;**205**(1–3):184–92.
107. Frake P, Greenhalgh D, Grierson SM, Hempenstall JM, Rudd DR. Process control and end-point determination of a fluid bed granulation by application of near infra-red spectroscopy. *Int J Pharm* 1997;**151**(1):75–80.
108. Landin M, York P, Cliff MJ, Rowe RC, Wigmore AJ. The effect of batch size on scale-up of a pharmaceutical granulation in a fixed bowl mixer granulator. *Int J Pharm* 1996;**134**(1–2):243–6.
109. Ohike A, Ashihara K, Ibuki R. Granulation monitoring by fast fourier transform technique. *Chem Pharm Bull* 1999;**47**(12):1734–9.
110. Watano S. Direct control of wet granulation processes by image processing system. *Powder Technol* 2001;**117**(1–2):163–72.
111. Whitaker M, Baker GR, Westrup J, Goulding PA, Rudd DR, Belchamber RM, et al. Application of acoustic emission to the monitoring and end point determination of a high shear granulation process. *Int J Pharm* 2000;**205**(1–2):79–91.
112. Lindberg N-O, Leander L, Reenstierna B. Instrumentation of a kenwood major domestic-type mixer for studies of granulation. *Drug Dev Ind Pharm* 1982;**8**(5):775–82.
113. Sakr WF, Ibrahim MA, Alanazi FK, Sakr AA. Upgrading wet granulation monitoring from hand squeeze test to mixing torque rheometry. *Saudi Pharm J* 2012;**20**(1):9–19.
114. Watano S, Okamoto T, Sato Y, Osako Y. Scale-up of high shear granulation based on the internal stress measurement. *Chem Pharm Bull* 2005;**53**(4):351–4.
115. Sato Y, Okamoto T, Watano S. Scale-up of high shear granulation based on agitation power. *Chem Pharm Bull* 2005;**53**(12):1547–50.
116. Hancock BC, York P, Rowe RC. Characterization of wet masses using a mixer torque rheometer: 2. Mixing kinetics. *Int J Pharm* 1982;**83**(1–3):147–53.
117. Hancock BC, York P, Rowe RC. An assessment of substrate-binder interactions in model wet masses. 1: mixer torque rheometry. *Int J Pharm* 1994;**102**(1–3):167–76.
118. Kristensen HG. Particle agglomeration in high shear mixers. *Powder Technol* 1996;**88**(3):197–202.
119. Corvari V, Fry W, Seibert W, Augsburger L. Instrumentation of a high-shear mixer: evaluation and comparison of a new capacitive sensor, a watt meter, and a strain-gage torque sensor for wet granulation monitoring. *Pharm Res* 1992;**9**(12):1525–33.
120. Arp Z, Smith B, Dycus E, O'Grady D. Optimization of a high shear wet granulation process using focused beam reflectance measurement and particle vision microscope technologies. *J Pharm Sci* 2011;**100**(8):3431–40.
121. Tok A, Goh X, Ng W, Tan RH. Monitoring granulation rate processes using three PAT tools in a pilot-scale fluidized bed. *AAPS PharmSciTech* 2008;**9**(4):1083–91.
122. Findlay WP, Peck GR, Morris KR. Determination of fluidized bed granulation end point using near-infrared spectroscopy and phenomenological analysis. *J Pharm Sci* 2005;**94**(3):604–12.
123. Briens L, Daniher D, Tallevi A. Monitoring high-shear granulation using sound and vibration measurements. *Int J Pharm* 2007;**331**(1):54–60.
124. Rantanen J, Wikström H, Turner R, Taylor LS. Use of in-line near-infrared spectroscopy in combination with chemometrics for improved understanding of pharmaceutical processes. *Anal Chem* 2004;**77**(2):556–63.
125. Li W, Worosila GD, Wang W, Mascaro T. Determination of polymorph conversion of an active pharmaceutical ingredient in wet granulation using NIR calibration models generated from the premix blends. *J Pharm Sci* 2005;**94**(12):2800–6.
126. Luukkonen P, Fransson M, Björn IN, Hautala J, Lagerholm B, Folestad S. Real-time assessment of granule and tablet properties using in-line data from a high-shear granulation process. *J Pharm Sci* 2008;**97**(2):950–9.
127. Narang AS, Sheverev VA, Stepaniuk V, Badawy S, Stevens T, Macias K, et al. Real-time assessment of granule densification in high shear wet granulation and application to scale-up of a placebo and a brivanib alaninate formulation. *J Pharm Sci* 2015;**104**(3):1019–34.
128. Luukkonen P, Schæfer T, Hellén L, Juppo AM, Yliruusi J. Rheological characterization of microcrystalline cellulose and silicified microcrystalline cellulose wet masses using a mixer torque rheometer. *Int J Pharm* 1999;**188**(2):181–92.
129. Hancock BC, York P, Rowe RC, Parker MD. Characterization of wet masses using a mixer torque rheometer: 1. Effect of instrument geometry. *Int J Pharm* 1991;**76**(3):239–45.
130. Parker M, Rowe R, Upjohn N. Mixer torque rheometry: a method for quantifying the consistency of wet granulations. *Pharm Tech Int* 1990;**2**(8):50–64.
131. Landín M, Rowe RC, York P. Characterization of wet powder masses with a mixer torque rheometer. 3. Nonlinear effects of shaft speed and sample weight. *J Pharm Sci* 1995;**84**(5):557–60.
132. Leuenberger H. Granulation, new techniques. *Pharm Acta Helv* 1982;**57**(3):72–82.
133. Sirois PJ, Craig GD. Scaleup of a high-shear granulation process using a normalized impeller work parameter. *Pharm Dev Technol* 2000;**5**(3):365–74.
134. Ghanta SR, Srinivas R, Rhodes CT. Use of mixer-torque measurements as an aid to optimizing wet granulation process. *Drug Dev Ind Pharm* 1984;**10**(2):305–11.
135. Faure A, Grimsey IM, Rowe RC, York P, Cliff MJ. Applicability of a scale-up methodology for wet granulation processes in collette gral high shear mixer-granulators. *Eur J Pharm Sci* 1999;**8**(2):85–93.
136. Landin M, York P, Cliff MJ, Rowe RC. Scaleup of a pharmaceutical granulation in planetary mixers. *Pharm Dev Technol* 1999;**4**(2):145–50.
137. Barrasso D, Ramachandran R. A comparison of model order reduction techniques for a four-dimensional population balance model describing multi-component wet granulation processes. *Chem Eng Sci* 2012;**80**(0):380–92.
138. Chaudhury A, Barrasso D, Pandey P, Wu H, Ramachandran R. Population balance model development, validation, and prediction of CQAs of a high-shear wet granulation process: towards QbD in drug product pharmaceutical manufacturing. *J Pharm Innov* 2014;**9**(1):53–64.

139. Darelius A, Rasmuson A, Björn IN, Folestad S. High shear wet granulation modelling—a mechanistic approach using population balances. *Powder Technol* 2005;**160**(3):209–18.
140. Gantt JA, Gatzke EP. High-shear granulation modeling using a discrete element simulation approach. *Powder Technol* 2005;**156**(2–3):195–212.
141. Immanuel CD, Doyle Iii FJ. Solution technique for a multi-dimensional population balance model describing granulation processes. *Powder Technol* 2005;**156**(2–3):213–25.
142. Kumar A, Gernaey KV, Beer TD, Nopens I. Model-based analysis of high shear wet granulation from batch to continuous processes in pharmaceutical production – a critical review. *Eur J Pharm Biopharm 85* 2013;(3, Part B):814–32.
143. Nakamura H, Fujii H, Watano S. Scale-up of high shear mixer-granulator based on discrete element analysis. *Powder Technol* 2013;**236**(0):149–56.
144. Poon JMH, Ramachandran R, Sanders CFW, Glaser T, Immanuel CD, Doyle Iii FJ, et al. Experimental validation studies on a multi-dimensional and multi-scale population balance model of batch granulation. *Chem Eng Sci* 2009;**64**(4):775–86.
145. Ramachandran R, Immanuel CD, Stepanek F, Litster JD, Doyle III FJ. A mechanistic model for breakage in population balances of granulation: theoretical kernel development and experimental validation. *Chem Eng Res Design* 2009;**87**(4):598–614.
146. Kayrak-Talay D, Dale S, Wassgren C, Litster J. Quality by design for wet granulation in pharmaceutical processing: assessing models for a priori design and scaling. *Powder Technol* 2013;**240**:7–18.
147. Iveson SM. Limitations of one-dimensional population balance models of wet granulation processes. *Powder Technol* 2002;**124**(3):219–29.
148. Annapragada A, Neilly J. On the modelling of granulation processes: a short note. *Powder Technol* 1996;**89**(1):83–4.
149. Adetayo AA, Ennis BJ. A new approach to modeling granulation processes for simulation and control purposes. *Powder Technol* 2000;**108**(2–3):202–9.
150. Gantt JA, Cameron IT, Litster JD, Gatzke EP. Determination of coalescence kernels for high-shear granulation using DEM simulations. *Powder Technol* 2006;**170**(2):53–63.
151. Liu LX, Litster JD. Population balance modelling of granulation with a physically based coalescence kernel. *Chem Eng Sci* 2002;**57**(12):2183–91.
152. Adetayo AA, Litster JD, Pratsinis SE, Ennis BJ. Population balance modelling of drum granulation of materials with wide size distribution. *Powder Technol* 1995;**82**(1):37–49.
153. Kapur PC. Kinetics of granulation by non-random coalescence mechanism. *Chem Eng Sci* 1972;**27**(10):1863–9.
154. Sastry KVS. Similarity size distribution of agglomerates during their growth by coalescence in granulation or green pelletization. *Int J Mineral Process* 1975;**2**(2):187–203.
155. Poon JMH, Immanuel CD, Doyle IFJ, Litster JD. A three-dimensional population balance model of granulation with a mechanistic representation of the nucleation and aggregation phenomena. *Chem Eng Sci* 2008;**63**(5):1315–29.
156. Tan HS, Goldschmidt MJV, Boerefijn R, Hounslow MJ, Salman AD, Kuipers JAM. Building population balance model for fluidized bed melt granulation: lessons from kinetic theory of granular flow. *Powder Technol* 2004;**142**(2–3):103–9.
157. Braumann A, Kraft M. Incorporating experimental uncertainties into multivariate granulation modelling. *Chem Eng Sci* 2010;**65**(3):1088–100.
158. Marshall CL, Rajniak P, Matsoukas T. Multi-component population balance modeling of granulation with continuous addition of binder. *Powder Technol* 2013;**236**:211–20.
159. Chaudhury A, Wu H, Khan M, Ramachandran R. A mechanistic population balance model for granulation processes: effect of process and formulation parameters. *Chem Eng Sci* 2014;**107**(0):76–92.
160. Remy B, Khinast JG, Glasser BJ. Discrete element simulation of free flowing grains in a four-bladed mixer. *AIChE J* 2009;**55**(8):2035–48.
161. Kuo HP, Knight PC, Parker DJ, Adams MJ, Seville JPK. Discrete element simulations of a high-shear mixer. *Adv Powder Technol* 2004;**15**(3):297–309.
162. Talu I, Tardos GI, Khan MI. Computer simulation of wet granulation. *Powder Technol* 2000;**110**(1–2):59–75.
163. Terashita K, Nishimura T, Natsuyama S. Optimization of operating conditions in a high-shear mixer using DEM model: determination of optimal fill level. *Chem Pharm Bull* 2002;**50**(12):1550–7.
164. Sato Y, Nakamura H, Watano S. Numerical analysis of agitation torque and particle motion in a high shear mixer. *Powder Technol* 2008;**186**(2):130–6.
165. Barrasso D, Tamrakar A, Ramachandran R. A reduced order PBM-ANN model of a multi-scale PBM-DEM description of a wet granulation process. *Chem Eng Sci* 2014;**119**:319–29.
166. Barrasso D, Eppinger T, Pereira FE, Aglave R, Debus K, Bermingham SK, et al. A multi-scale, mechanistic model of a wet granulation process using a novel bi-directional PBM-DEM coupling algorithm. *Chem Eng Sci* 2015;**123**:500–13.
167. Barrasso D, Ramachandran R. Multi-scale modeling of granulation processes: Bi-directional coupling of PBM with DEM via collision frequencies. *Chem Eng Res Design* 2015;**93**:304–17.

CHAPTER

29

Process Development, Optimization, and Scale-Up: Fluid-Bed Granulation

K. Yamamoto[1] *and Z.J. Shao*[2]

[1]Pfizer Global Research and Development, Groton, CT, United States
[2]Arena Pharmaceuticals Inc., San Diego, CA, United States

29.1 OVERVIEW OF THE FLUID-BED GRANULATION PROCESS

One commonly used strategy to improve undesirable powder characteristics of the active pharmaceutical ingredient (API) is to granulate the API with other excipients. There are two widely used methods to make granules: dry granulation, which involves mechanical compaction (slugging or roller compaction) followed by a dry-sizing process; and wet granulation using a liquid binder (low shear wet granulation, high shear wet granulation, or fluid-bed granulation). While separate drying equipment is required following a low-shear or high-shear wet granulation process, fluid-bed wet granulation is a one-pot process that can blend, wet-granulate, and dry, all in one piece of equipment.[1]

A typical fluid-bed wet granulation sequence consists of dry blending, wet granulation, and drying steps. Raw materials are first fluidized by hot inlet air for blending. Granulation fluid is then sprayed onto the preblended material until the desired moisture content or granule size is achieved. At this point, drying begins until a predetermined product temperature is reached. Binder solution is commonly used as a granulation fluid. However, it can also be added as a dry ingredient to the powder bed. Granules produced using the fluid-bed process have the advantages of being homogeneous, free-flowing, and less dense. Generally, these granules require lower pressure to compress into tablets than granules made by other granulation methods.

The amount of solvent removed during fluid-bed wet granulation and drying steps can be calculated using heat and mass transfer theory. By taking process variables, such as spray rate, dew point, volume of inlet air, and powder characteristics of raw material into consideration, a robust manufacturing process can be developed for a variety of formulation compositions and manufacturing scales.

Being a self-enclosed system, the fluid-bed is a preferred technology for highly potent compounds requiring containment. When coupled with a clean-in-place (CIP) system, fluid-bed becomes the preferred granulation method for such potent compounds.

29.2 EQUIPMENT DESIGN

29.2.1 Batchwise models

Among the three well-established batch fluid-bed methods, top spraying is commonly used for granulating purposes due to its ability to produce highly porous particles with better dispersibility and compressibility for downstream compression. The bottom-spray design (Wurster) is very popular for layering of active component and coating to modify drug release due to its good film formation in such units. Drug-layering on sugar beads using Wurster inserts is widely reported in the research literature. Tangential spray (rotary) methods are mostly used in pelletization and coating of the formed pellets. The operating mechanisms are shown in Fig. 29.1. The design and operational details for the three methods are described in the next sections.

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00029-7

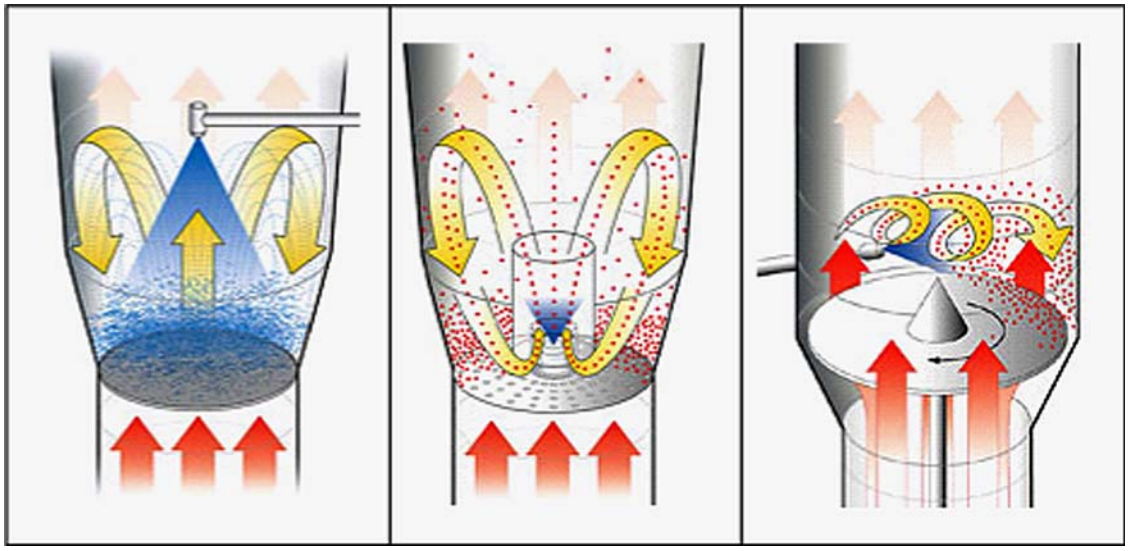


FIGURE 29.1 Top spray, bottom spray, and tangential spray mechanisms (from left to right) used in batch fluid-beds. *Source: Courtesy Glatt Air Techniques, Inc.*

29.2.1.1 Top spray

The main components for a top-spray processor are:

- An air-handling unit optionally equipped with humidification and dehumidification for dew-point control
- A product container
- An expansion chamber
- A spray system containing single or multiple nozzles
- An exhaust system with a filter and, optionally, an explosion suppression system

Glatt expansion chambers come in models known as tall conical GPCG design or the shorter and wider WSG design. GPCG machines are suitable for top-spray and bottom-spray granulation, while the WSG design is suited for drying of wet granules, and occasionally for top-spray granulation processes.

Top spray has a simple setup, offers a high capacity, and is typically used for either granulating powders or spray-drying. When a binder solution is sprayed through the nozzles, the droplets produced bind to fluidizing powder particles, forming agglomerates. There are cases when water is sprayed into the powder bed, which contains a dry binder, to form granules. Granulation with water, also referred to as *instantizing*, is mostly used in the food and diagnostic industries. Granules formed by instantizing are generally more loosely formed and friable. Controlling inlet air dew point and bed temperature is important for batch-to-batch consistency.

A top-spray fluid-bed can also be used as an alternative to the traditional spray-drying process using hot air. Since the air volume in a fluid-bed can be increased easily, it does not demand a high temperature to dry-sprayed droplets. This is especially attractive to protein formulations, which can be denatured if the processing temperature is too high. Particles formed through a top-spray fluid-bed method are larger in size due to the effect of layering. Flowability of these particles is generally better, as a result of larger particle formation with narrow size distribution.

Since the particles formed are very porous, the top-spray method is not suitable for coating of beads and fine particles for controlled-release products. One exception is hot-melt coating, where molten materials are sprayed onto fluidized particles. Jones and Percel[2] used all three methods to produce sustained-release chlorpheniramine maleate pellets by spraying molten, partially hydrogenated cottonseed oil, and found that the top-spray method resulted in a smooth surface and a decreased dissolution rate.

29.2.1.2 Bottom spray

Bottom spraying, originally developed by Dale Wurster in the late 1950s, is most commonly used in the coating of particulates and tablets. There are several unique features with a Wurster-type column design. First, the nozzle is mounted at the bottom of the product container and above the air distribution plate. Second, the air distribution plate has larger holes in the middle than on the outer rim. This allows more air flow through the middle portion of the plate. Finally, an insert (partition) is placed above the plate. This configuration ensures that particles are fluidized upward in the insert, fall outside the insert, pass under

the insert, and are lifted again inside the insert. The region inside the insert is commonly referred to as the *up-bed region*, while that outside the insert is known as the *down-bed region*. Particles are sprayed in the up-bed and dried in the down-bed many times, until the desired film thickness is achieved.

Since coating solution is sprayed from the bottom, coating efficiency is typically much higher than by top spraying. Mehta and Jones[3] demonstrated that bottom spraying produced a much smoother coating than top spraying, especially when organic solvent-based solution was sprayed. Top-spraying solvent-based solution leads to flash evaporation and more spray-drying, resulting in a rough surface.

29.2.1.3 *Tangential spray*

The tangential-spray process involves a variable-speed rotating disk inside the product bowl. The spray nozzle is mounted at the side of the bowl near the bottom and is buried in the batch during processing. The gap between the disk and bowl wall can be adjusted to allow varying amounts of air through. During granulation, the rotating disk imparts a centrifugal force to the material, forcing the particles to move to the outer part of the disk. The air flow then provides a lifting force. These two forces, along with natural gravity, make the material travel in a spiraling helix. This motion results in the compaction of the material to produce uniform and dense granules, thereby enabling the manufacture of high-dose loading products. An added advantage is that further coating of the formed pellets can be done in the same processor, without having to discharge the pellets. Disk speed and air volume need to be reduced, compared to the granulation stage, in order to avoid film breakage potentially caused by abrasion.

29.2.2 Semicontinuous design

The best known case of implemented semicontinuous fluid-bed technology is the Glatt MultiCell,[4,5] which combines high shear granulation and fluid-bed drying (Fig. 29.2). Typically, three fluid-bed dryers are set up side by side, along with a horizontal high shear granulator and a wet sieving mill. After wet granulation, the wet mass is wet-screened and loaded in the first fluid-bed for partial drying (predrying). Partially dried granules are then pneumatically conveyed to the second fluid-bed unit for continued drying (final

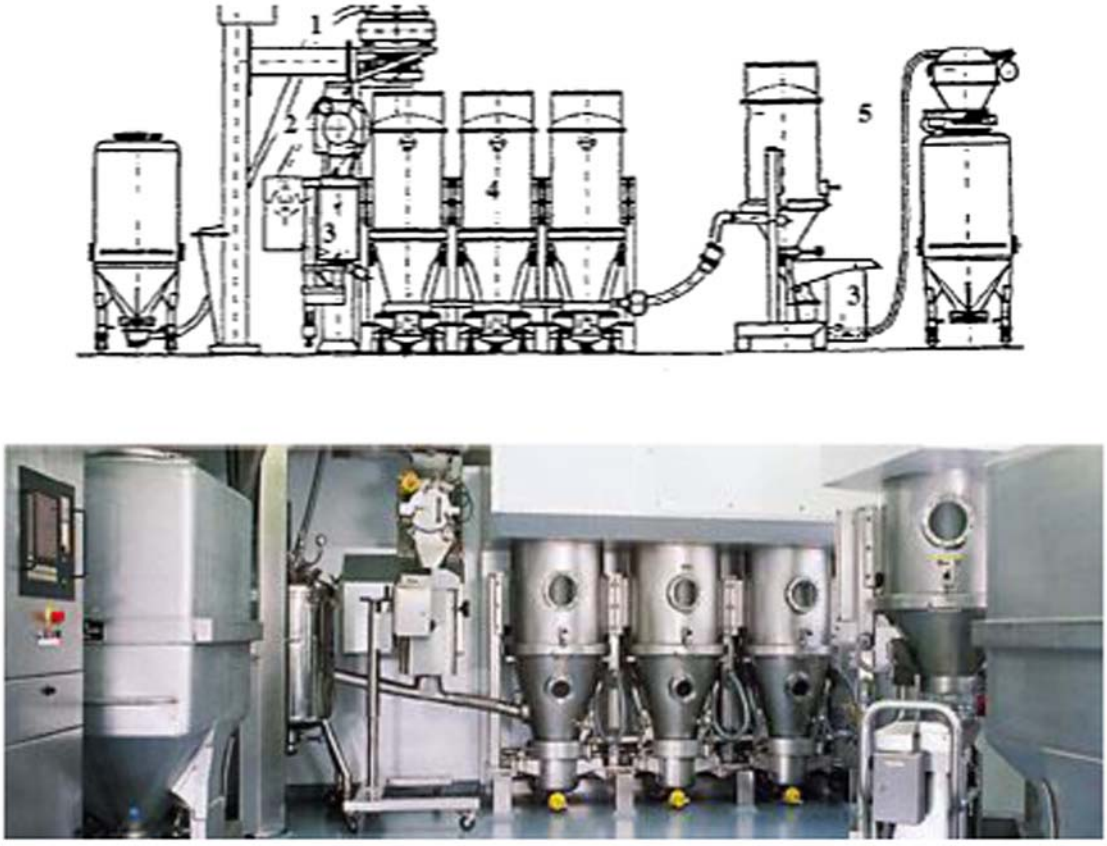


FIGURE 29.2 Semicontinuous fluid-beds, Glatt multicell: 1. Dosing unit; 2. high-shear granulator; 3. sieve; 4. fluid-bed; 5. pneumatic transport system. *Source: Courtesy Glatt Air Techniques, Inc.*

drying), then to the third unit for final drying and equilibration (cooling). The dried granules are then conveyed to a dry mill for milling.

The three fluid-bed units are designed for different air temperature and humidity conditions, with the first one typically set at high temperature (60°C), and the last cell at ambient temperature and humidity.

GEA Niro Pharma Systems (Columbia, MD) demonstrated a new version of a semicontinuous fluid-bed coater a few years ago, the SUPERCELL.[6,7] This device is capable of automatically preweighing small batches of tablets (typically 30–120 g) into the coating chamber, coating them rapidly (typically <10 minutes), and finally pneumatically conveying them out of the chamber. The ability of SUPERCELL to coat tablets and capsules rapidly and uniformly is claimed to be due to uniquely patented designs (ie, a special gas distribution plate), and a two-fluid spray nozzle located below it.

The atomizing gas is mixed with low-pressure drying gas, thereby reducing the momentum of the atomizing air. Attenuation of the spraying momentum, along with a very short processing time, is expected to reduce tablet wear and tear, thereby improving tablet appearance. Conceivably, friable or hygroscopic tablets can also be processed in such a unit due to its rather short coating time. A vendor-sponsored study of coating an enteric film on aspirin tablets confirmed good acid resistance for tablets coated with >9% weight gain. No tablet edge defects after coating were found. A prototype model is shown in Fig. 29.3. Acceptance and use of this device by the pharmaceutical industry has yet to be seen.

29.2.3 Continuous models

Glatt developed various continuous fluid-bed models with the GF and AGT units shown in Fig. 29.4. In the GF model, the inlet air plenum is divided into multiple chambers, each with adjustable temperature and air flow control. By this means, spray agglomeration, drying, and cooling can all be done in this single unit. By controlling air flow in the various chambers, the overall powder flows in a plug fashion toward the discharge port.

In the 1980s, Glatt commercialized the AGT technology, which has a discharge pipe right in the middle of the bottom screen. The air velocity in this pipe determines what size particles can be discharged. Heavy particles with sinking velocity greater than the air velocity inside the pipe will be discharged, while smaller particles are blown back to the fluid-bed and are exposed to layering again.

FIGURE 29.3 Cutaway view of the liquid nozzle and gas distribution plate, tablet movement inside the coating chamber, and the SUPERCELL coater general appearance. *Source: Courtesy* GEA Niro Pharma Systems.

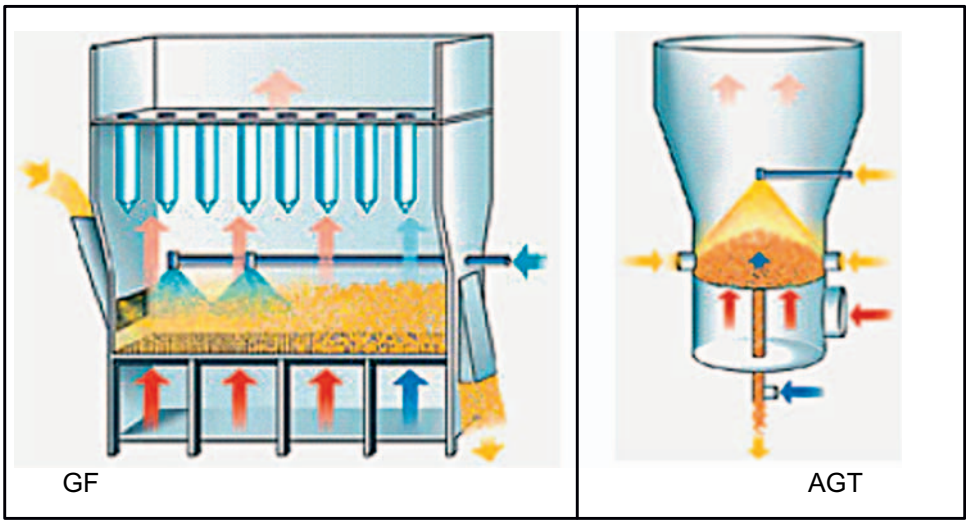


FIGURE 29.4 Continuous fluid-beds, Glatt GF and AGT units. *Source: Courtesy Glatt Air Techniques, Inc.*

Continuous processors are typically used in the chemical, food, and agricultural industries.

29.3 FLUID-BED HYDRODYNAMICS

29.3.1 Product temperature and moisture content profiles through fluid-bed processing

A fluid-bed, wet-granulation process consists of dry blending, wet granulation, and drying. A typical example of product temperature and moisture content profiles through fluid-bed processing is that product temperature initially increases during the blending step. It then decreases after spraying starts and eventually bottoms out. Assuming that there is no heat loss from the equipment surface, and that the materials are well mixed by fluidization air, product temperature could reach:

- Wet bulb temperature of the inlet air when the spray rate is greater than the drying capacity of the inlet air. In this case, product temperature is kept constant at the wet bulb temperature, even after drying begins. This continues until water levels in the wet mass become less than the drying capacity, at which time the product temperature rises.
- Between the inlet air temperature and the wet bulb temperature of the inlet air, if the spray rate is lower than the drying capacity of the inlet air. In this case, the product temperature starts increasing immediately after spraying is stopped.

Once product temperature starts rising, a good correlation between the moisture content of granules and product temperature can be established. Product temperature becomes a good indicator for predicting the drying end point.

Following moisture level changes through fluid-bed processing yields more information than monitoring product temperature. The moisture level of starting materials first decreases during the blending step, and then it increases when spraying starts and keeps increasing until spraying stops. The accumulation of moisture depends on processing parameters (inlet air and spraying conditions). A good correlation between moisture content and size of wet granules can also be established. When the spraying step is completed, accumulated moisture within the wet mass starts to be removed. It has also been demonstrated that a good correlation between moisture content and size of dry granules exists.

29.3.2 Moisture mass balance during the fluid-bed process

Moisture profiling throughout the fluid-bed process not only can be used as a fingerprint for the formulation and process, but is also useful for process development and troubleshooting purposes. When wet granules contact the hot inlet air, the hot inlet air exchanges its heat for moisture. The temperature of the air drops due to evaporative cooling and exits the process chamber as exhaust air. Water in wet granules is vaporized instantly by heat at this temperature. Since the granules are suspended into fluidization (inlet) air, and contact surface area between air and wet granules is considered to be huge, heat and mass transfer reach equilibrium instantaneously.

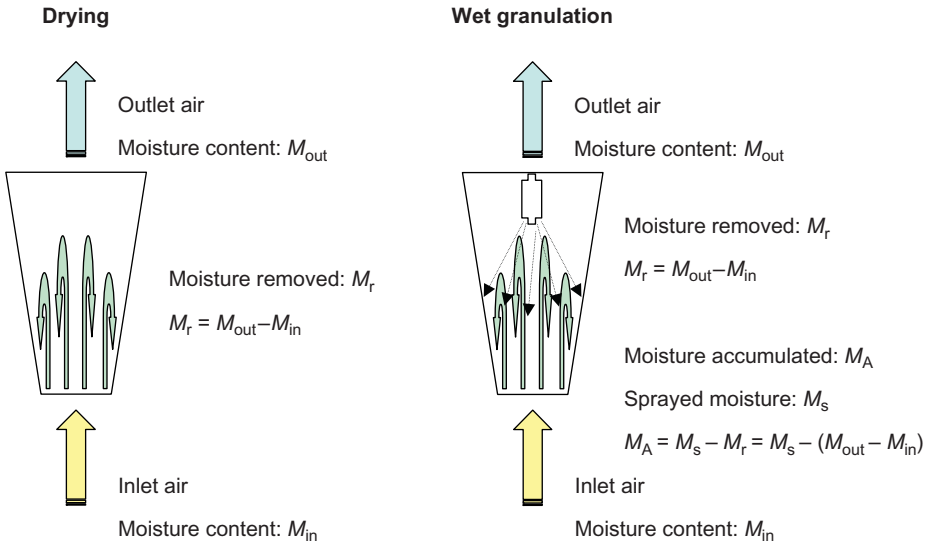


FIGURE 29.5 Moisture mass balance during fluid-bed process.

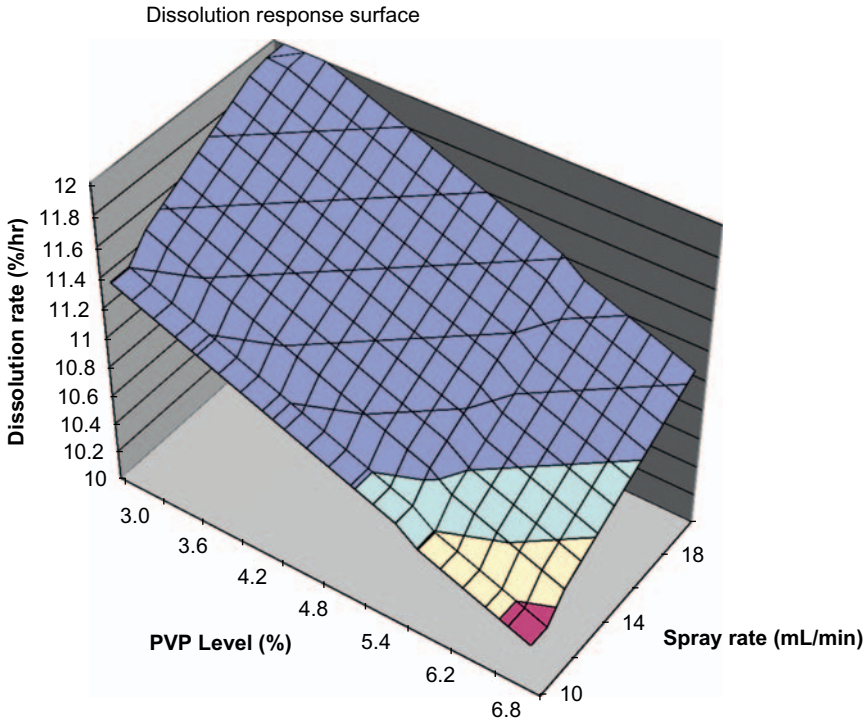


FIGURE 29.6 Dissolution response surface from a DOE.

29.3.2.1 *Drying process*

Fig. 29.5 shows moisture balance during fluid-bed operation. In the drying process, moisture in wet granules is removed by dry inlet air. During wet granulation, there is another source of moisture, which is sprayed binder solution.

As shown in Fig. 29.6, moisture removed (M_r) during the drying process can be calculated from the moisture content of inlet (M_{in}) and outlet air (M_{out}) as follows:

$$M_r = M_{out} - M_{in}$$

where M_{in} and M_{out} can be determined by monitoring the temperature and relative humidity of inlet and outlet air. Using moisture content (M_{end}) and weight (W_{ini}) of wet granules at the wet-granulation end

point, the moisture content of wet granules during the drying process can be calculated:

$$M_g = \frac{(M_{end} - M_r)}{W_{ini} - M_r} \times 100$$

29.3.2.2 Wet-granulation process

To predict moisture balance during the wet-granulation process, two parameters need to be taken into consideration: moisture removed (M_r), and moisture accumulated in the wet granules (M_A) for the wet-granulation process (Fig. 29.5). Since the binder solution is sprayed onto the powder bed during this process, moisture in the sprayed binder solution (M_s) needs to be built into the equation in order to calculate the moisture level accumulated in wet granules:

$$M_A = M_s - M_r = M_s - (M_{out} - M_{in})$$

The moisture content of wet granules can be calculated using the initial conditions, which include the batch size (W_{ini}), and initial moisture content of the powder bed (M_{ini}):

$$M_g = \frac{M_{ini} + M_A}{W_{ini} + M_A} \times 100$$

29.3.2.3 Prediction of moisture profile during the fluid-bed process

By taking air factors, including air volume per unit time and specific volume, into consideration, the moisture profile during fluid-bed operation can be predicted. Briefly, a given inlet air drying capacity is expressed as

$$DC = H_{in\_wb} - H_{in\_db}$$

where H_{in_wb} and H_{in_db} are the absolute moisture content of the wet bulb and the dry bulb temperature of inlet air, respectively.

The rate of moisture removal (X_r) is expressed as

$$X_r = \left(H_{in\_wb} \times \frac{V_{in}}{VH_{sp\_out}}\right) - \left(H_{in\_db} \times \frac{V_{in}}{VH_{sp\_in}}\right)$$

where V_{in} is inlet air volume per unit time, and VH_{sp_in} and VH_{sp_out} are specific volume of the inlet air and outlet air, respectively.

Then the moisture content for the drying process is expressed as

$$M_g = \frac{(M_{end} - X_r \times t)}{W_{ini} - X_r \times t} \times 100$$

and the moisture content for the wet-granulation process is expressed as

$$M_g = \frac{M_{ini} + (X_s - X_r) \times t}{W_{ini} + (X_s - X_r) \times t} \times 100$$

where X_s and t are spray rate and running time, respectively.

29.4 MECHANISMS OF AGGLOMERATION

29.4.1 Phases in granule growth

The main objective of granulation is to improve the undesirable powder characteristics of raw materials (poor powder flow, fluffiness, segregation, etc.); therefore, the specific needs first must be identified. Granulation properties after the process are likely different, depending on the target of the process.

The moisture content of wet granules can be controlled by process parameters. If there is a target moisture range in mind, it can generally be achieved through process controls (air and spray conditions). At the beginning of the process, the powder bed is assumed to be dry, with no aggregation. Following binder spraying, the powder bed becomes wet, and then aggregation starts and granule size increases with the spraying of more binder solution. The granule growth rate during the wet-granulation process can vary from slow progress in the beginning, to rapid growth exceeding the desired end point. To develop a robust wet-granulation process, the basic concept of granule growth and its influencing factors need to be understood.

29.4.2 Bonding mechanisms

Rumph[8] classified particle–particle bonding mechanisms into five categories:

1. Solid bridges
2. Interfacial forces and capillary pressure in movable liquid surfaces
3. Adhesion and cohesion forces in bonding bridges which are not freely movable
4. Attraction between solid particles, such as van der Waals forces, electrostatic forces, and magnetic type interactions
5. Form-closed bonds, such as particles interlocking or entangling

Category 1 is the moisture-mediated bonding mechanism, with interfacial and capillary forces between particles playing important roles to form particle–particle bonds during the wet granulation process. Particle–particle bonds are classified into four categories, based on the phase of interaction between liquid and particles: pendular, funicular, capillary, and droplet, according to Newitt and Conway-Jones and Barlow.[9,10] Assuming that a binder solution is used as the granulation fluid, which is sprayed into fine droplets, liquid bridges between particles can be viscous and adhesive. Formed solid bridges will remain after drying.

Upon spraying binder solution onto the powder bed, moisture content increases and the surface of powders becomes self-adhesive due to coalescence of the binder solution. During the coalescence, water may be absorbed into particles, resulting in high void saturation with increased deformation capability (plasticity) which helps to bind particles together (see Huang and Kono,[11,12] and Abberger[13]). Once liquid bridges start to be formed between particles (nucleation), the progress of granule growth is determined by a balance of cohesive and adhesive forces between particles (viscosity of the liquid bridge), and impact from collision of particles during the wet-granulation process.[14] Since two or more particles are held together with liquid bridges, granule size could be regulated by number, size, and viscosity of the liquid bridges. If the liquid bridge size is small or less viscous, then the bridge can hold only small particles or breakage may occur. If the liquid bridge size is large, viscous, or both, the liquid bridges could hold heavier and larger particles, as well as granules, together.

29.5 FORMULATION AND PROCESS VARIABLES AND THEIR CONTROL

Like many other unit operations, many variables can affect manufacturability and product quality in the fluid-bed wet-granulation process. In order to develop a robust process, it is necessary to evaluate the effects of these variables at an early stage. These variables can be classified into three groups: equipment-related, formulation-related, and process-related variables. The performance of a fluid-bed wet granulator depends on equipment-related variables, such as equipment size (dimensions) and shape, position of spray nozzle (top, bottom, tangential), number of spray guns, etc. These variables need to be assessed during the scale-up activity from one scale to another. This aspect was covered in Section 29.6, earlier in this chapter.

Due to the basic mechanism of granule growth during the wet-granulation process, particle size and its distribution, density, wettability, solubility, and hygroscopisity of starting materials impact product quality (bioperformance, content uniformity, chemical stability, etc.), and manufacturability of the formulation. It is also well known that cohesiveness and static charge of raw material can cause undesirable homogeneity and low-yield problems after the wet-granulation process. On the other hand, inlet air conditions (relative humidity of dew point, temperature, air volume, or velocity), and spraying conditions (spraying rate, atomizing air pressure, spray gun position), are key variables to make a consistent quality product batch after batch (reproducibility).

29.5.1 Formulation variables

There are a variety of excipients commercially available for fluid-bed formulation development, and each excipient has different material properties. It is important to understand excipient functionality and its behavior during processing. Since multiple excipients are put into a formulation, it then becomes complicated to understand the impact of all these excipients on manufacturability and the final product quality.

Schaefer and Worts investigated the influences of starting material property and its mixing ratio in the fluid-bed granulation process, using crystalline lactose and starch as raw materials. It was shown that the granule size depended on the ratio of lactose and starch, with higher lactose content generating larger granule sizes. It was also indicated that the higher the surface area of starting materials, the smaller the granule size that was obtained after granulation.[15] Wan et al. also demonstrated that, after fluid-bed wet granulation, the granule size of lactose, a water-soluble excipient, is larger than the granule size of cornstarch, a water-insoluble excipient.[16] Since different excipients have different particle sizes and interact differently with water, processing conditions will need to be different in order to obtain similar granule properties.

A binder is an essential component of the granulation process, which provides binding ability in the formulation after drying. The binder can be added either as powder or sprayed as a solution to the powder bed during the wet-granulation process. Wan and Lim reported that larger granule size and higher density is obtained when the binder is added as a powder, rather than by spraying as a solution.[16] Schaefer and Worts demonstrated that the droplet size of the binder solution is influenced by the viscosity of the binder solution. However, it was also shown that viscosity is not the sole factor that affects granule size.[17] The type of binders and the binder concentration could be critical variables when the binder solution is sprayed using a nozzle system. Binders such as hydroxypropylcellulose, polyvinylpyrollidone (PVP), hydroxypropylmethylcellulose (HPMC), and pregelatinized starch are commonly used for fluid-bed wet granulation with a range of 2–5% w/w solution.[18]

29.5.2 Key process variables

As discussed in the previous section, there are three major factors influencing the fluid-bed wet granulation process: namely, air conditions, spray conditions, and equipment-related factors. The equipment-related factors include design of equipment, bag filter, distributor, shape of bowl and extension chamber, shaking mechanisms, and the position and number of spray

nozzles. These factors need to be carefully evaluated when batch size is changed or a formulation is transferred from one site to another.

29.5.2.1 Inlet air conditions

Air conditions, including temperature, dew point (or absolute humidity), and inlet air volume, are the three key factors influencing drying and wet granulation. Since moisture from wet granules is evaporated and removed as exhaust air during the wet granulation and drying process, exhaust air contains a higher amount of moisture than inlet air. As described in Section 29.3.2 earlier in this chapter, the amount of moisture that inlet air can remove (*DC*) is

$$DC = H_{\text{in\_wb}} - H_{\text{in\_db}}$$

Absolute moisture values depend on inlet air, which is the reason why the dew point of inlet air and inlet air temperature are the critical factors for the wet-granulation and drying processes. The same inlet air temperature and inlet air volume could result in a different *DC* value due to a different dew point. If no dew point controller or dew point monitoring system is available, it is recommended to measure inlet air relative humidity and temperature to make the *DC* consistent (manufacturing process reproducible). These absolute moisture values can be obtained from the psychometric cart.

The powder bed is fluidized by inlet air, and it is important to keep sufficient fluidization throughout the process; otherwise, the bed may collapse. It may be necessary to adjust inlet air volume during the wet-granulation and drying processes because of weight change of wet versus dry materials. Inlet air volume needs to be reduced in order to minimize attrition during the drying process. The amount of moisture removed per unit time is expressed as

$$\frac{\text{Amount of moisture removed}}{\text{minute}} = DC \times \text{air volume/minute}$$

When inlet air volume is adjusted, the amount of moisture removed per unit time will change.

29.5.2.2 Spray rate, droplet size, and spray pattern

Since formation of liquid bridges among particles is the main mechanism of granule growth, it is reasonable to expect that spray rate and droplet size significantly affect the wet-granulation process and influence granule size and size distribution. The effects of spray rate and droplet size on fluid-bed wet granulation are not much different from those for the high-shear wet-granulation process. If the spray rate is too fast, rapid granule growth occurs, and the process could be sensitive to subtle changes in raw material and process conditions. Water-soluble components may be dissolved into the binder solution, which accelerates granule growth. If the spray rate is too slow, it takes a long time to reach the desirable end point, and in the worst cases, no granule growth occurs. Large droplet size tends to produce larger and denser granules, while a fine mist may result in spray-drying the binder solution. Spray pattern and atomizing air influence droplet size. Shaefer and Worts[17,18] reported that a broad spray pattern creates insufficient air and liquid mixing, which causes the droplet size to become larger (Table 29.1). Small changes of air and liquid flow ratio (air-to-liquid mass ratio) influence droplet size changes (Table 29.2).

The effect of droplet size and powder material particle size on granule size after fluid-bed melt granulation using polyethylene glycol 3000 was examined by Abberger et al. [19] It was concluded that the mechanism of agglomeration depended on the ratio of the binder droplet size to the particle size of the solid powder. When the solid powder particle size is smaller than the binder droplet size, immersion of the solid particle in the surface of the binder solution is the main mechanism of nucleation. Distribution of the binder solution to the solid particle makes the solid particle's surface sticky and deformable. When the solid particle size is greater than the binder droplet size, these sticky and deformable wet particles stick together, which becomes the dominant nucleation mechanism.

TABLE 29.1 Influence of Spray Angle on Droplet Size (d_{50}, μm)

Air dome setting		
2	5	1.2
(c. 30 degree)	(c. 40 degree)	(c. 60 degree)
56 μm	89 μm	100 μm

Data from Schaefer T, Worts O. Control of fluidized bed granulation I. Effects of spray angle, nozzle height and starting materials on granule size and size distribution. Arch Pharm Chem Sci Ed *1977;5:51–60; binder solution: gelatin 4%, 40°C; liquid flow rate: 150 g/min; air flow rate: 10 Nm³/h.*

TABLE 29.2 Influence of Liquid Flow Rate and Air-to-Liquid Mass Ratio on Droplet Size (d_{50}, μm)

	Liquid flow rate (g/min)		
Mass ratio	100	150	200
1.15	129 μm	99 μm	98 μm
	123 μm	113 μm	98 μm
1.43	106 μm	90 μm	79 μm
	102 μm	86 μm	88 μm

Data from Schaefer T, Worts O. Control of fluidized bed granulation I. Effects of spray angle, nozzle height and starting materials on granule size and size distribution. Arch Pharm Chem Sci Ed *1977;5:51–60; binder solution: gelatin 4%, 40°C.*

29.5.3 Granule growth under drier conditions (low moisture content of wet granules during the granulation process)

AbuBaker et al.[20] investigated the effects of processing conditions on granule properties using a $2^{5-1}+2$ factorial design of experiment (DOE), using a fluid-bed granulator/dryer. The factors evaluated were fine and coarse drug particle size, binder solution concentration (11% and 16%), spray rate (70 and 110 g/minutes), atomizing air pressure (1 and 3 bar), and inlet air dew point (-5°C and 15°C). Statistical analysis from the study indicated that the spray rate and atomization air pressure ratio significantly influence granule size. Due to the very low moisture content (less than 2%) of granules, distribution of binder solution to solid particle surface makes the particle surface ready for aggregation, which is likely to be the main mechanism of granule growth. In this case, factors such as the spray rate and atomization pressure (liquid-to-air) ratio, which controls the droplet size of binder solution, become dominant factors for granule growth. The level of moisture content in wet granules may not be a good parameter to monitor, due to the very small change of the moisture level during the process.

29.5.4 Granule growth under wetter conditions (high moisture content of wet granules)

When moisture is accumulated in wet granules, more liquid bridges are formed, which increases the contact surface area between particles. As a result, the bonding capability of wet particles becomes higher than that for drier particles. It was reported that granule size is proportional to the humidity level inside the fluid-bed,[21] so granulation size can be monitored by measuring the moisture content of the wet mass. Watano et al.[22,23,24] introduced an infrared (IR) sensor to monitor the moisture level during fluid-bed granulation and established a good correlation between moisture content and granule size.

29.6 SCALE-UP CONSIDERATIONS

Scale-up of fluid-bed processes from small laboratory units to large commercial machines has been a continuing activity in the chemical and pharmaceutical industries for over half a century. In spite of this, fluid-bed scale-up is still not an exact science; it remains a mix of mathematics, engineering, and personal judgment. There are many simplifications, approximations, and educated guesses involved in fluid-bed granulation and drying. Better data, more realistic models, and more exact equations are always sought after in this field. The secret to successful fluid-bed scale-up, however, lies in the recognition and management of uncertainties rather than finding the fluid-bed's exact attributes.[25] Generally, equipment variables, such as the type and size of the equipment, and key process variables, such as spray rate, atomization pressure, and inlet air temperature, affect the product quality attributes. Control of such parameters, in order to yield a consistent product at a large batch size (scale), thereby constitutes a successful scale-up strategy. Equally important is the consistent quality of incoming raw materials (API and excipients). A change in particle size distribution of the API is known to affect granulation characteristics, and ultimately compressibility.

Of the three steps involved in fluid-bed agglomeration (dry mixing, spray agglomeration, and drying), the spray agglomeration stage is the most critical phase to monitor. During this phase, dynamic granule growth and breakdown takes place, along with solvent evaporation. Schaefer and Worts[21] demonstrated that granule size is directly proportional to bed humidity during granulation. Therefore, controlling bed humidity becomes critical. Gore et al.[26] studied various processing parameters and concluded that air temperature, spray nozzle location, spray rate, and atomization pressure can all affect granule characteristics. If one has to limit the number of process parameters to control for scale-up purposes, bed humidity and atomizing air pressure should be chosen. Atomization pressure affects granule growth through its effect on droplet size and spray pattern. Maintaining binder droplet size, when performing scale-up activities, is regarded as the key to scale-up success. Along with atomization pressure, nozzle atomization air volume should be recorded.

The primary consideration in scaling up the fluid-bed process is to maintain drying efficiency between the laboratory and production equipment. In a larger unit, inlet air temperature may need to be lowered, due to its higher air volume, in order to maintain the same drying efficiency as in smaller units. Otherwise, spray-drying of binder solution can occur. Another commonly used method is to maintain bed temperature, which can be achieved by balancing air temperature, dew point, air flow, and spray rate.

An additional consideration during scale-up is material handling procedures and transfer methods. While hand-scooping into and out of the bowl is convenient for small-scale processors, it is physically demanding and may not even be acceptable from a worker safety perspective. Vacuum transfer becomes a preferred option and a common practice. This is achieved through the creation of negative pressure inside the unit by running the blower with the inlet air flap minimally open and leaving the outlet flap fully open. Ingredients can be charged one by one via vacuum, and then the mixing and granulation phases can be started. For granule discharge, the

product can be vacuum transferred, bottom discharged by gravity, or side discharged into another intermediate bulk container (IBC) for further mixing and compression. Fully contained IBCs aided by vacuum transfer that can be linked to the discharge port of the fluid-bed processor are commercially available. These units keep the bowl from opening during product discharge and are typically used for potent compounds. If the fluid-bed is equipped with CIP cleaning nozzles, it can then be washed down before the bowl disengages from the expansion chamber.

29.6.1 Batch size and equipment selection

Scale-up from small laboratory sized fluid-bed machines can be made much easier if the same line of equipment is used. Manufacturers generally keep scale-up issues in mind when designing their line of processors. However, if equipment from a different manufacturer is to be used at the commercial site due to availability limitations, efforts will need to be spent on modifying process parameters. This is because of differences in air flow pattern, expansion chamber geometry, gun-spray pattern, and other factors.

Minimum and maximum batch size within selected top-spraying equipment can be approximated using the following equations:

$$S_{\min} = V \times 0.5 \times BD$$

$$S_{\max} = V \times 0.8 \times BD$$

where S is batch size in kilograms, V is the product bowl working volume in liters, and BD is the bulk density of finished granules in grams per cubic centimeter.

Linear scale-up can be used as a rough estimate in the beginning. Assuming that 6-kg scale batches have been made in a 22-L bowl at 60% volume, and that the next scale-up will be done in a 100-L bowl, a starting batch size of 30 kg is recommended. One must be kept in mind that as the batch size increases in the larger units, the powder bed will become more compacted, resulting in a higher bulk density in the bowl. Jones[27] predicted that approximately 20% higher bulk density should be factored in when comparing larger units to smaller ones.

For tangential-spray in a rotary fluid-bed granulator, minimum and maximum batch sizes can be determined as follows:

$$S_{\min} = V \times 0.2 \times BD$$

$$S_{\max} = V \times 0.8 \times BD$$

Rotor granulation is amenable to a wider range of batch size changes in the same bowl, although one should keep in mind that a larger batch load typically produces denser granulation due to the increased mass effect. Similar to a top-spraying model, a good starting point would be for the finished product to occupy 60% of the working capacity or volume of the rotor product chamber.

For bottom-spraying in a Wurster-type of equipment, the minimum and maximum batch size can be calculated using the following expressions:

$$S_{\min} = \frac{1}{2}(\pi R_1^2 H - N\pi R_2^2 H) \times BD$$

$$S_{\max} = (\pi R_1^2 H - N\pi R_2^2 H) \times BD$$

where R_1 is the radius of the chamber, R_2 is the radius of the partition, N is the number of partitions, and H is the length of the partition.

29.6.2 Spray rate scale-up

Spray rate scale-up is determined by the drying capacity of the equipment rather than by the increase in batch size. In other words, spray rate change should be based on similarities in fluidization patterns at different scales. However, one should keep in mind that at a given atomization pressure and air flow volume, change in liquid spray rate affects droplet size, which in turn affects particle agglomeration.

Two methods can be used to determine spray rates. If air volume can be read directly from the equipment, spray rate can be simply calculated as follows:

$$S_2 = S_1 \times \frac{V_2}{V_1}$$

where S_1 is spray rate in the laboratory scale equipment, S_2 is spray rate in the scaled-up equipment, V_1 is air volume in the laboratory scale equipment, and V_2 is air volume in the scaled-up equipment.

If there is no direct air volume readout on the equipment, cross-sectional areas of the product bowl screens can be used for approximation as follows:

$$S_2 = S_1 \times \frac{A_2}{A_1}$$

where A_1 is the cross-sectional area of the laboratory scale equipment, and A_2 is the cross-sectional area of the scaled-up equipment.

Using a 22-L bowl as the starting point, scale-up and scale-down factors have been tabulated for ease of use, as listed in Table 29.3.

Recently, AbuBaker et al.[20] proposed that the ratio of spray rate/atomization pressure is the most important factor controlling the droplet size of the spray. The authors reached this conclusion while working on scaling up an eroding tablet formulation with a high

TABLE 29.3 Approximate Scale-Up Factors for Fluid-Bed Granulators

Unit volume (L)	Screen diameter (mm)	Spray rate scale-up factor	Run-time scale-up factor
1.75	100	0.21	0.38
4.50	150	0.47	0.44
22	220	1.0	1.00
45	350	2.5	0.82
100	500	5.2	0.87
215	730	11.0	0.89
420	900	16.7	1.14
670	1000	20.7	1.47
1020	1150	27.3	1.70
1560	1250	32.3	2.20
2200	1750	63.3	1.58
3000	1740	62.6	2.18

drug load. By maintaining a constant ratio, scale-up from 8 to 70 kg was successful despite differences in drug particle size, granulator geometry, and number of nozzles used.

The authors have further devised a new process parameter termed *driving force*, which is the product of spray rate, binder concentration, and droplet size, and then normalized by batch size. Granule growth rate was shown to correlate well with calculated driving force:

$$\text{Driving Force} = \frac{\text{Spray Rate} \times \text{Binder Concentration} \times \text{Droplet Size}}{\text{Batch Size}}$$

Conceivably, if driving force is maintained constant at different scales, scale-up can be done with ease, resulting in granules with the same particle size.

29.6.3 Rotary disk speed scale-up

A simple way to scale-up disk speed is by maintaining radial velocity (V) constant as the diameter of the disk increases:

$$N_2 = N_1 \times \frac{d_1}{d_2}$$

where N is the disk speed (rpm) and d is the diameter of the disk (m), and the subscripts 1 and 2 represent the first and second set of conditions, respectively.

Radial velocity (V, m/s) can be calculated from

$$V = \frac{\pi d N}{60}$$

TABLE 29.4 Glatt RSG and GRG Rotor Parameters

Model	Disk diameter	Radial velocity (m/s)	Speed (rpm)	Air volume (m^3/h)	Number of nozzles
RSG 1/3	295	0–27	0–1800	80–580	1–2
RSG 5/15	485	0–30	0–1200	120–750	1–3
RSG 15/30	620	3.7–22	113–680	200–1500	2–4
RSG 60/100	780	4.4–22	108–540	600–3000	2–4
GRG 100	1000	5–25	100–500	800–4500	3–6
GRG 200	1400	5–25	70–340	1000–6000	3–6
GRG 300	1600	5–25	60–300	1400–8000	4–8

Alternatively, radial acceleration can be held constant in order to keep steady the centrifugal force (F_c) acting on the bed. Since radial acceleration equals to $2 \times V^2/d$, the following scale-change equation can be derived:

$$N_2 = \sqrt{\frac{N_1^2 d_1}{d_2}}$$

Assuming that a disk speed of 200 rpm was used in a 485-mm rotor, the scaled-up disk speed for the 780-mm rotor would be

$$\text{Using radial velocity method: } N_2 = \frac{200 \times 0.485}{0.780} = 124 \text{ rpm}$$

Using radial acceleration method:

$$N_2 = \sqrt{\frac{(200)^2 \times 0.485}{0.780}} = 158 \text{ rpm}$$

The known equipment parameters for the Glatt RSG and GRG rotor granulators are listed in Table 29.4.

Scale-up in a rotor granulator can also be done by maintaining F_c constant for different-sized rotors or the same rotor with different amounts of material loading. Horsthuis et al.[28] used the Froude number for scale-up in high shear Gral compounders. Chukwumezie et al.[29] modified the Froude number method to maintain F_c in rotor granulation. By maintaining an F_c of 41,667 N for a 1-kg batch in a 12-in. rotor at 500 rpm, these authors were able to scale the process up to a 10-kg load in a 19-in. rotor at 200 rpm:

$$F_c = \frac{W \times V^2}{R}$$

where W is the batch weight, V is the radial velocity, and R is the plate radius.

TABLE 29.5 CPCG, RSG, and GRG Batch Size Ranges

GPCG	RSG	GRG	Disk (mm)	Maximum working volume (L)	Typical batch size range (kg)
1 & 3	1	N/A	300	4.5	0.5–3.0
5 & 15	5	N/A	500	30	3–20
30 & 60	30	30	620	60	5–40
60 & 120	60	60	780	105	10–75
N/A	N/A	100	1000	180	25–125
N/A	N/A	200	1400	430	50–300

Typical batch size ranges for the GPCG, RSG, and GRG lines of machines are listed in Table 29.5.

29.6.4 Rational scale-up

Due to the large numbers of interrelating parameters affecting product attributes, scale-up efforts can be significantly improved by employing rational approaches, such as DOE. DOE allows the scientist to narrow down the most critical formulation and process parameters to monitor during scale-up trials. Full factorial designs afford the identification of factor-to-factor interactions that cannot be revealed by the one-parameter-at-a-time method.

As an example, a controlled-release formulation was granulated in a small-scale GPCG-3 top-spray granulator in the authors' laboratory. A $2^3 + 2$ design (three factors, two levels, with two center points) was utilized in this study, with the factors being PVP level, spray rate, and HPMC viscosity variation within the same commercial grade. Resulting granulations were compressed into tablets and tested for dissolution. The response surface plot is shown in Fig. 29.6. While there were no significant effects of spray rate and HPMC viscosity on dissolution, the PVP level did appear to be a statistically significant factor. Therefore, controlling PVP levels in the formulation is important, although how PVP affects HPMC gelation during dissolution still remains unclear.

Process analytical technology (PAT) also proves to be a valuable tool in process monitoring. Among well-established PAT methodologies is the near-infrared (NIR) spectroscopy technique via fiber-optic probes placed inside the bowl of the granulator. Since a wide spectrum (1100–2500 nm) of signals is captured in real time along the granulation process, it can then be processed to reflect water and particle size changes. Frake et al.[30] installed an inline NIR system in a Glatt GPCG 30/60. The second derivative changes in absorbance at 1932 nm were used to correlate with moisture, while the zero-order absorbance at 2282 nm was selected to reflect particle size increase during granulation. This continual inline NIR monitoring has been shown to provide a suitable means of end-point determination for a top-spray granulation process.

Rantanen et al.[31] employed a multichannel NIR detector in a Glatt WSG-5 unit to monitor moisture level during granulation and drying of blends and pellets. Moisture detection was performed at 1990 nm throughout the mixing, spraying, and drying phases. Together with bed temperature measurements, NIR moisture monitoring provides a value-added tool for better process understanding.

Watano et al.[32] utilized an IR moisture sensor for a laser scattering spray size analyzer for their work with agitation fluid-beds. They recommended that if the ratio of spray area to vessel cross-sectional area is kept constant, localized overwetting can be prevented. By means of moisture control, drying efficiency during wetting and agglomeration phases can be maintained constant among the different sized vessels. In a follow-up article,[33] these researchers further correlated the relationship between IR absorbance and granule moisture content. Except for very dry conditions, such as low spray rate or high air temperature, there appears to be excellent correlation between measured moisture level (by offline drying) and IR absorbance.

Another method in monitoring real-time granule growth is by image analysis. Watano and Miyanami[34] developed an image processing system by coupling a particle image probe (ie, a charge-coupled device camera), and an image-processing unit. A close agreement in granule size between image process and the traditional sieve analysis was obtained. The authors further automated this system to make it capable of controlling the spray rate at the end of the granulation phase or in order to avoid excessive agglomerate formation.[35] Modulation of spray rate was accomplished via fuzzy logic, a linguistic algorithm employing if-then rules, and a process lag element. Good control of granule growth was reported.

29.6.5 Scale-up via semicontinuous (batch-continuous) processing

The principle of using continuous processing to improve cost efficiency is well established in the food and bulk chemical industries. The pharmaceutical industry has been rather slow to adopt such a strategy due to relatively high profit margins, constant product changeover at the manufacturing site, and equipment qualification concerns. Semicontinuous processes, however, offer several advantages over traditional batchwise processing. The most important of these is

the ability to manufacture small and large volumes in the same equipment setup, thereby eliminating scale-up needs. Other advantages include ease of automation and less product risk in case of batch failure.[36]

Betz et al.,[37] using the Glatt MultiCell and two placebo formulations, demonstrated that consistent yield (discharge from the high shear granulator), compression profiles, and tablet disintegration can be generated throughout many subunits manufactured. This method has been extensively tested by Roche Basel, with more than 30 marketed products. It was postulated that 600 subunits can be manufactured to constitute a large "batch" of 4200 kg.[37] One other such subunit set-up has been installed in Pfizer's facility in Freiburg, Germany.

29.6.6 Scale-up via continuous processing

True continuous fluid-bed processing equipment is rarely used by the pharmaceutical industry, although suitable equipment lines are available. The primary applications are the chemical and food industries. An example is the Glatt GF series, which were developed over a decade ago in order to improve the throughput of fluid-bed processes. Originally designed for drying solids, various models are available today for granulation, coating, pelletizing, drying, and cooling applications.

The design feature of the GFG (granulator) is its inlet air plenum, which is divided into multiple chambers, thereby enabling the introduction of air with different temperatures and velocity in each chamber. Together with the correct placement of spray nozzles above each chamber, it is possible to create very different conditions in different sections of the process chamber. Agglomeration, drying, and cooling, therefore, can be done as the product travels from chamber to chamber within the same unit.

29.7 APPLICATION OF QUALITY-BY-DESIGN TO FLUID-BED GRANULATION

The inter-related nature of fluid-bed granulation process parameters can be approached more effectively via quality-by-design principles, especially with the aid of PAT. Critical quality attributes (CQAs) defined during granulation and drying steps typically encompass granulation particle size distribution, moisture content, bulk density, tapped density, and flowability index. Through the appropriate use of DOE, data analysis, and modeling, critical process parameters can be identified and suitable design space defined in the regulatory filing for operational flexibility. It is perfectly feasible to envision real-time batch release of granules from granulators upon successful implantation of online or inline PAT techniques which are validated to provide instantaneous data on CQAs. Drying can be terminated using real-time data and built-in feedback loops in the batch manufacturing recipe.

Peinado et al.[38] described a case study on the establishment of a validated NIR method to monitor granulation moisture level at commercial scale and compared with data generated using traditional loss on drying (LOD) methodology. A 300-L Glatt fluid-bed dryer was instrumented with an inline NIR spectrophotometer and signals captured between 1178 and 2075 nm, upon which a partial least squares model was then built. After proving that applicable validation criteria (specificity, linearity, accuracy, robustness, and precision) were all met, this method has been approved as the primary drying end-point control for this product, yielding much improved precision over the conventional LOD method.

Lourenço et al.[39] reported another design space case study on a large scale (WSG200 drier, 500 kg) by leveraging an inline moisture analyzer based on microwave resonance technology and another inline particle size analyzer based on spatial filter velocimetry (SFV). Samples were also characterized offline using optical microscopy, laser diffraction, and Karl-Fisher titration. A partial factorial DOE was executed to evaluate inlet air temperature, air flow rate, and binder spray rate during the spray phase. Inlet air temperature and air flow rate during drying were also included as process parameters in this DOE. Granulation moisture content has been shown to be affected by inlet air temperature, binder-spray rate, and air flow rate increment (recipe changes) during the granulation phase. Based on this study, a design space for the commercial scale granulation process has been proposed (which happens to be within previously filed limits).

In addition to NIR and SFV, there are other technologies that are capable of determining granulation particle size inline (ie, focused beam reflectance measurements, acoustic emission, particle image velocimetry, etc.). These techniques, along with data interpretation through statistical treatments,[40] can help in delineating critical process parameters and their multivariate interactions, which ultimately define a design space in regulatory applications.

29.8 SUMMARY

Fluid-bed granulation is a widely used and well-characterized unit operation in the pharmaceutical industry. With proper selection of equipment design, operating conditions, and suitable excipients, it has been shown that this technology can be scaled up from

the laboratory to commercial production. Although the wet-granulation process can vary from formulation to formulation, it is important to understand the basic concepts of equipment design and fundamental granulation theory, in order to develop robust and scalable formulations and processes. Application of PAT and DOE can help establish a design space for the fluid-bed process, as well as serving as great tools for scale-up and technology transfer. When a formulation and process are optimally developed using fluid-bed technology, the granules will ultimately impart superb quality to the end product (tablet, capsule, etc.).

References

1. Rubino OP. Fluid-bed technology. *Pharm Technol* 1999;**23**(6):104–13.
2. Jones DM, Percel PJ. Coating of multiparticulates using molten materials: formulation and process considerations. In: Ghebre-Sellassie, I, editor. *Multiparticulate oral drug delivery*, Vol. I. New York, NY: Marcel Dekker; 1994. p. 113–42.
3. Mehta AM, Jones DM. Coated pellets under the microscope. *Pharm Technol* 1985;**9**(6):52–60.
4. Leuenberger H. New trends in the production of pharmaceutical granules: batch versus continuous processing. *Eur J Pharm Biopharm* 2001;**52**:289–96.
5. Betz G, Junker-Burgin P, Leuenberger H. Batch and continuous processing in the production of pharmaceutical granules. *Pharm Dev Technol* 2003;**8**:289–97.
6. Birkmire, A.P., Walter, K.T., Liew, C.V. & Tang, E.S.K. (2004). Tablet coating in the novel SUPERCELL™ coater: evaluation of color uniformity. Poster presented at the 2004 AAPS annual meeting and exposition, Baltimore, MD.
7. Felton, L., Sturtevant, S. & Birkmire, A. (2006). A novel capsule coating process for the application of enteric coatings to small batch sizes. Poster presented at the 2006 AAPS annual meeting and exposition, San Antonio, TX.
8. Rumpf H. *Agglomeration*. New York, NY: John Wiley & Sons Inc; 1962. p. 379.
9. Newitt DM, Conway-Jones JM. A contribution to the theory and practice of granulation. *Trans Inst Chem Eng* 1958;**36**:422–42.
10. Barlow CG. Granulation of powders. *Chem Eng* 1968;**220**:CE196–201.
11. Huang CC, Kono HO. The granulation of partially prewetted alumina powders—a new concept in coalescence mechanism. *Powder Technol* 1988;**55**:19–34.
12. Huang CC, Kono HO. A mathematical coalescence mode in the batch fluidized bed granulator. *Powder Technol* 1988;**55**:35–49.
13. Abberger T. The effect of powder type, free moisture and deformation behaviour of granules on the kinetics of fluid-bed granulation. *Eur J Pharm Biopharm* 2001;**52**:327–36.
14. Ennis BJ, Tardos G, Pfeffer R. A microlevel-based characterization pf granulation phenomena. *Powder Technol* 1991;**65**:257–72.
15. Schaefer T, Worts O. Control of fluidized bed granulation IV. Effects of binder solution and atomization on granule size and size distribution. *Arch Pharm Chem Sci Ed* 1977;**5**:14–25.
16. Wan LSC, Lim KS. Mode of action of polyvinylpyrrolidone as a binder on fluidized bed granulation of lactose and starch granules. *STP Pharma Prat* 1989;**5**:244–50.
17. Schaefer T, Worts O. Control of fluidized bed granulation I. Effects of spray angle, nozzle height and starting materials on granule size and size distribution. *Arch Pharm Chem Sci Ed* 1977;**5**:51–60.
18. Schaefer T, Worts O. Control of fluidized bed granulation II. Estimation of droplet size of atomized binder solutions. *Arch Pharm Chem Sci Ed* 1977;**5**:178–93.
19. Abberger T, Seo A, Schaefer T. The effect of droplet size and powder particle size on the mechanism of nucleation and growth in fluid bed melt agglomeration. *Int J Pharm* 2002;**249**:185–97.
20. AbuBaker O, Canter K, Ghosh S, Hedden DB, Kott L, Pipkorn D, et al. Development of a novel scale up parameter to optimize and predict fluid bed granulation using experimental design (DOE). *AAPS PharmSci* 2003;**5**(4). Abstract T3295.
21. Shaefer T, Worts O. Control of fluidized bed granulation III. Effect of inlet air temperature and liquid flow rate on granule size and size distribution. Control of moisture content of granules in the drying phase. *Arch Pharm Chem Sci Ed* 1977;**6**:1–13.
22. Watano S, Fukushima T, Miyanami K. Heat transfer and granule growth rate in fluidized bed granulation. *Chem Pharm Bull (Tokyo)* 1996;**44**:572–6.
23. Watabe S, Morikawa T, Miyanami K. Methematical model in the kinetics of agitation fluidized bed granulation. Effects of moisture content, damping speed and operation time on granule growth rate. *Chem Pharm Bull (Tokyo)* 1996;**44**:409–15.
24. Watabe, Takahashi H, Sato Y, Yasutomo T, Miyanami K. Measurement of moisture content by IR sensor in fluidized bed granulation. Effects of operating variables on the relationship between granule moisture content and absorbance of IR spectra. *Chem Pharm Bull (Tokyo)* 1996;**44**:1267–9.
25. Matsen JM. Scale-up of fluidized bed processes: principle and practice. *Powder Technol* 1996;**88**:237–44.
26. Gore AY, McFarland DW, Batuyios NH. Fluid bed granulation: factors affecting the process in laboratory development and production scale-up. *J Pharm Technol* 1985;**9**:114–22.
27. Jones DM. Factors to consider in fluid bed processing. *J Pharm Technol* 1985;**9**:50–62.
28. Horsthuis GJB, van Laarhoven JAH, van Rooij RCBM, Vromans H. Studies on upscaling parameters of the Gral high shear granulation process. *Int J Pharm* 1993;**92**:143–50.
29. Chukwumezie BN, Wojcik M, Malak P, Adeyeye MC. Feasibility studies in spheronization and scale-up of ibuprofen microparticulates using the rotor disk fluid-bed technology. *AAPS PharmSciTech* 2002;**3**(article 2):1–13.
30. Frake P, Greenhalgh D, Grierson SM, Hempenstall JM, Rudd DR. Process control and end-point determination of a fluid bed granulation by application of near infra-red spectroscopy. *Int J Pharm* 1997;**151**:75–80.
31. Rantanen J, Lehtola S, Rämet P, Mannermaa J-P, Yliruusi J. On-line monitoring of moisture content in an instrumented fluidized bed granulator with a multi-channel NIR moisture sensor. *Powder Technol* 1998;**99**:163–70.
32. Watano S, Sato Y, Miyanami K, Miyakami T, Ito T, Kamata T, et al. Scale-up of agitation fluidized bed granulation. I. Preliminary experimental approach for optimization of process variables. *Chem Pharm Bull (Tokyo)* 1995;**43**:1212–16.
33. Watano S, Takashima H, Sato Y, Yasutomo T, Miyanami K. Measurement of moisture content by IR sensor in fluidized bed granulation. Effects of operating variables on the relationship between granule moisture content and absorbance of IR spectra. *Chem Pharm Bull (Tokyo)* 1996;**44**:1267–9.
34. Watano S, Miyanami K. Image processing for on-line monitoring of granule size distribution and shape in fluidized bed granulation. *Powder Technol* 1995;**83**:55–60.

35. Watano S, Sato Y, Miyanami K. Control of granule growth in fluidized bed granulation by an image processing system. *Chem Pharm Bull (Tokyo)* 1996;**44**:1556–60.
36. Vervaet C, Remon JP. Continuous granulation in the pharmaceutical industry. *Chem Eng Sci* 2005;**60**:3949–57.
37. Betz G, Junker-Bürgin P, Leuenberger H. Batch and continuous processing in the production of pharmaceutical granules. *Pharm Dev Technol* 2003;**8**:289–97.
38. Peinado A, Hammond J, Scott A. Development, validation and transfer of a near infrared method to determine in-line the end point of a fluidised drying process for commercial production batches of an approved oral solid dose pharmaceutical product. *J Pharm Biomed Anal* 2011;**54**:13–20.
39. Lourenço V, Lochmann D, Reich G, Menezes JC, Herdling T, Schewitz J. A quality by design study applied to an industrial pharmaceutical fluid bed granulation. *Eur J Pharm Biopharm* 2012;**81**:438–47.
40. Tomba E, Facco P, Bezzo F, Barolo M. Latent variable modeling to assist the implementation of quality-by-design paradigms in pharmaceutical development and manufacturing: a review. *Int J Pharm* 2013;**457**:283–97.

CHAPTER

30

Formulation, Process Development, and Scale-Up: Spray-Drying Amorphous Solid Dispersions for Insoluble Drugs

B. Vig[1] *and M. Morgen*[2]

[1]Bristol-Myers Squibb Company, New Brunswick, NJ, United States [2]Bend Research, a division of Capsugel Dosage Form Solutions, Bend, OR, United States

30.1 INTRODUCTION

A majority of new active pharmaceutical ingredients (APIs) in development exhibit low aqueous solubility and poor dissolution limiting oral bioavailability. A number of approaches have been used to improve aqueous solubility and bioavailability.[1–4] These have included particle size reduction[5]; creation of amorphous solid dispersions (ASDs)[6,7]; use of solubilization formulations (eg, using complexing agents, cosolvents, and lipid-based formulations)[8,9]; crystal modification (eg, using salts and cocrystals)[10]; and chemical modification (eg, prodrugs).[11]

Amorphous solids offer a number of advantages for the delivery of APIs (also referred to as compounds or drugs) with low aqueous solubility. Amorphous solids promote faster dissolution and higher dissolved drug concentrations because they lack long-range order, exhibit weak interactions between individual molecules, and have higher enthalpy, entropy, and free energy than their crystalline counterparts.[12] However, some formulations based on amorphous solids may present challenges, including physical instability, poor manufacturability at a large scale, inconsistent performance, and materials handling difficulties. This chapter is focused on the use of ASDs, which overcome a number of these issues by achieving a kinetically stabilized amorphous form.[7] ASDs, which are molecular dispersions of drug, polymer(s), and/or other excipient(s), can be prepared using a number of processes and are applicable to a wide range of API types, offering an attractive means to increase the aqueous solubility and bioavailability of problematic compounds. This chapter addresses the formulation, process development, scale-up, and drug product considerations for preparing spray dried dispersions (SDDs), a particular class of ASDs, for APIs with low aqueous solubility.

30.2 BACKGROUND

In this section, we provide background information on ASDs in general, and SDDs in particular.

30.2.1 Background: assessment of ASD applicability

ASDs are increasingly used commercially and clinically because of their utility.[13,14] Successful use of ASDs, however, depends on careful assessment of the suitability of this approach for the specific application. Following this assessment, care must also be taken in the selection of the dosage form, taking into account performance (solubility enhancement and sustainment for biologically relevant timeframes), chemical and physical stability, manufacturability, and adequate drug loading, as described in the next sections.

The low aqueous solubility of compounds is generally attributed to strong interactions in the crystal lattice (high melting point) and/or hydrophobic

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00030-3

groups (as indicated by a high octanol/water partition coefficient, log *P*).[15] Insight into the utility of a particular solubilization application can be gained by assessing the dose-to-solubility ratio required to achieve high oral bioavailability for the compound. This can be done using a technology map, such as that shown in Fig. 30.1, which shows the applicability of ASDs versus other technologies. The top-most solid diagonal line in the map (Fig. 30.1) traces the maximal solubility (S_{max}) of the lowest-energy, neutral form of the compound, calculated using a modified general solubility equation (Eq. (30.1)). Eq. (30.1) assumes that compound solid-state interactions are negligible (ie, the compound is a liquid at ambient temperature).

$$S_{max}(\text{mg/mL}) = 1000 \times 10^{(-\text{Log } P)} \qquad (30.1)$$

Decreasing aqueous solubility at a constant log *P* value is driven primarily by an increase in the solid-state interactions, which are directly proportional to melting temperature (T_m). Generally, the further a compound falls below the diagonal line, the higher its T_m value. In the upper region of this map, crystalline solubility is high enough to achieve good bioavailability at a 100-mg active (mgA) dose using a crystalline drug form. Decreasing solubility associated with increasing log *P* and/or increasing T_m requires a solubilization-enabling technology to maintain good in vivo performance.

As Fig. 30.1 shows, particle-size-reduction technologies (eg, creating microcrystals or nanocrystals via micronization or milling) can offer acceptable bioavailability at a 100-mgA dose when solubility falls below 1 mg/mL. This is achieved by increasing the surface area to overcome the slow dissolution rate of a crystalline drug, keeping the drug at its equilibrium concentration in the gastrointestinal (GI) tract.[5,16]

However, as solubility decreases further, the utility of such technologies diminishes because absorption is inadequate even if high (including instantaneous) dissolution rates are achieved. At such low solubilities, technologies are required that increase the drug concentration in the GI lumen above its equilibrium solubility and/or that increase drug transport across the unstirred water layer via submicron-sized colloids. For compounds with high lipophilicity (ie, log $P > 6$), the addition of lipidic excipients can help solubilize and enhance transport of the compound through the unstirred aqueous boundary layer—a process that can otherwise be slow and limit absorption of lipophilic drugs.

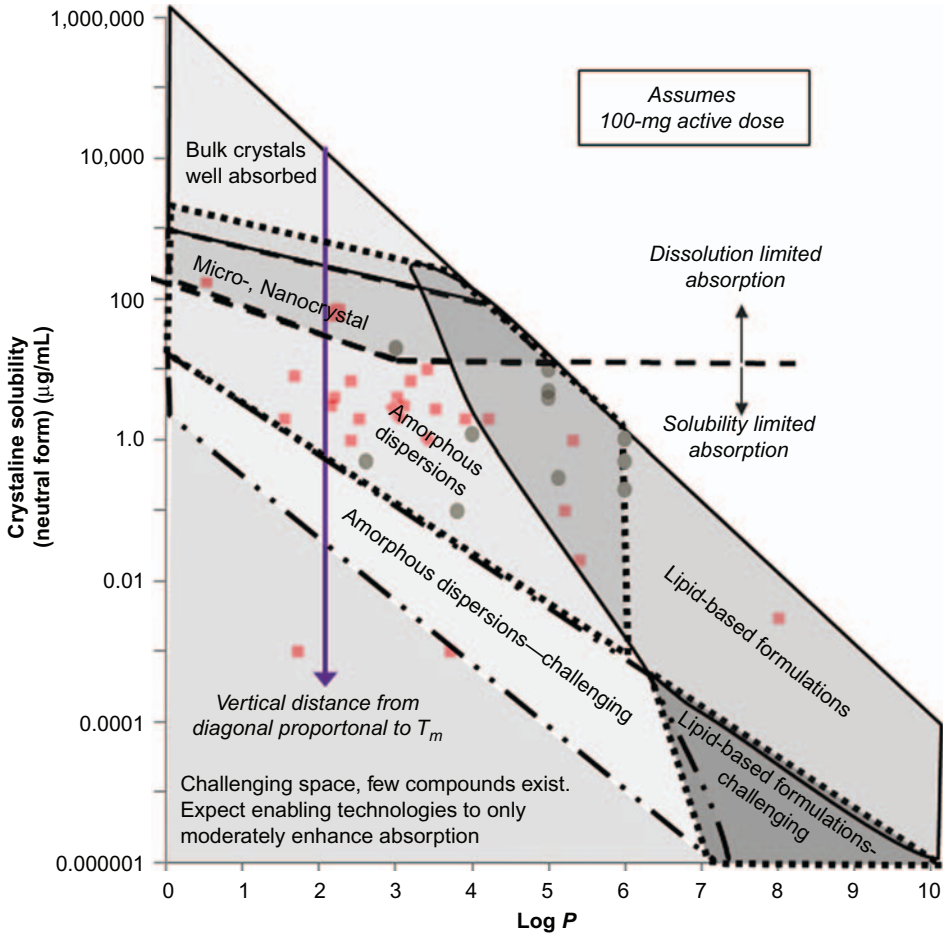


FIGURE 30.1 Crystalline solubility versus Log *P* of compounds showing applicability of Amorphous Solid Dispersions (ASDs) compared with other technologies. Assumptions: the figure is based on solubility of the neutral bulk crystalline form at pH 6.5, assuming the fraction of dose ionized at pH 7 is less than 50%, the molecular weight of the compound is less than approximately 700 Da (so permeability is not adversely affected by molecular size), and the human dose is 100 mgA. If the dose is higher or lower than this, the property space in which high bioavailability can be achieved changes.

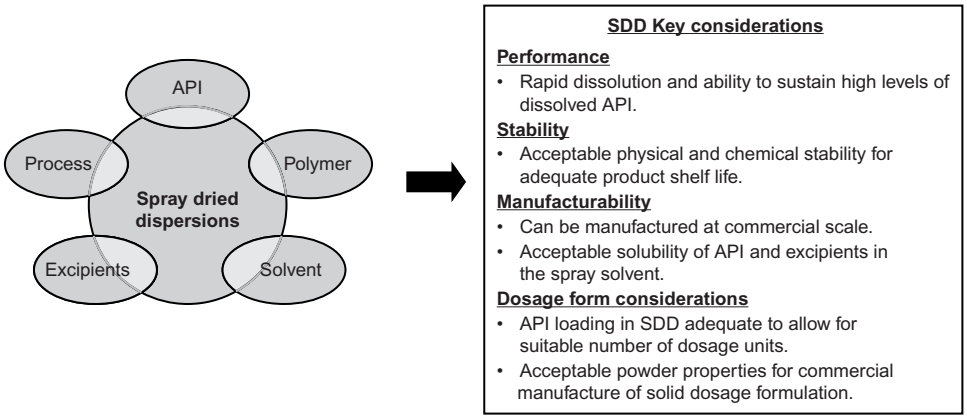


FIGURE 30.2 Spray Dried Dispersions (SDDs) Target Product Profile—Performance, stability, manufacturability, and dosage-form considerations.

Due to their rapid dissolution rate and higher drug activity relative to bulk crystalline drug, ASDs are useful where the drug at the desired dose is dissolution- or solubility-limited, often with a dose number >1.[17] ASDs are most readily applied to compounds that have low to moderate log *P* values, although the technology can be extended to compounds with higher log *P* values.[13,18] ASDs have been shown to keep dissolved drug levels high and enhance bioavailability significantly in preclinical and clinical settings.[19,20]

Typically, ASDs are used instead of pure amorphous API because the physical stability of the latter is usually not sufficient to avoid rapid API crystallization in the solid state or in the use environment. ASDs can be made by a number of processes, including spray-drying,[21] spray coating of API and excipient onto substrates,[22] solvent-controlled precipitation,[23] freeze drying,[24] fusion-based process,[25] and hot melt extrusion (HME).[26] Processing temperatures, miscibility of API and solvent, time spent in a high-mobility state, particle properties (eg, size and density), and equipment availability can all affect the choice of the process for making ASDs. For example, an API with a high T_m can present challenges for preparation using HME due to the need for higher processing temperatures that, in turn, increases the risk of API and/or excipient degradation.[26] This chapter is focused on ASDs made using spray-drying—one of the most common processes for converting API and excipients (typically polymers) into an ASD. The SDD shares some, but not all characteristics of ASDs made using other processes.

30.2.2 SDD key considerations

SDDs are prepared by spray drying a solution of API, polymer, and possibly other excipients, in a volatile spray solvent. SDD development, like other formulation development, should begin with a well thought out target product profile (TPP), which can guide the development process and be refined as additional information becomes available.[27,28] Ideally, SDD formulations for bioavailability enhancement should meet several manufacturing and performance criteria, as shown in Fig. 30.2. On the manufacturing side, the API and excipients should have relatively high solubility in one or more high-volatility spray solvents and have good chemical stability in solution. The solvent should be easy to remove during spray-drying and/or subsequent secondary drying. On the performance side, the formulation typically should have a rapid dissolution rate and sustain high dissolved-drug concentrations. Other key attributes include good physical stability with respect to amorphous phase separation and drug crystallization, and good solid-state chemical stability of the drug and excipients. Ideally, all these criteria should be met using a SDD with a high enough drug loading to meet the TPP, allowing delivery of the desired dose using an acceptable number of dosage units. Finally, SDDs should have acceptable powder properties for commercial manufacture of solid dosage formulation (eg, capsules or tablets). Table 30.1 lists common critical to quality attributes (CQAs) that can impact TPP and the formulation and process parameters that commonly define the SDD attributes.

30.3 SDD FORMULATION COMPOSITION

As shown in Fig. 30.2 and Table 30.1, SDD CQAs and, thus, performance is dependent on the formulation composition (eg, API and polymer chemistry, API:polymer ratio, and other excipients); choice of spray solvent(s); and the manufacturing process parameters (eg, droplet size, drying rate, and secondary drying). It is important to understand the interdependencies between the formulation composition and manufacturing process to ensure manufacture of a consistent SDD product that meets the desired performance criteria.

TABLE 30.1 Potential SDD Critical to Quality Attributes and Important Formulation and Process Parameters and Their Potential Product Impact

Potential SDD CQA	Potential product impact	SDD composition considerations	SDD process parameters consideration	Typical analytical tools
Physical state (presence of crystalline API)	• Product physical stability (presence of crystalline API) • Dissolution, in vivo performance	• API physicochemical properties eg, T_g, T_m, Log P, counterion, API-polymer miscibility • API to polymer ratio • Polymer chemistry and MW • Spray solvent • Spray solution solids content	• Droplet size and drying rate • Incomplete dissolution of API in spray solvent. • Secondary drying parameters • Exposure to humidity	• DSC • PXRD • SEM and PLM • FT-IR • Raman • ssNMR • Dissolution tests
Potency	• Potency	• API to polymer ratio • Spray solution solids content	• Incomplete API and/or polymer dissolution in spray solvent	• HPLC • In-line spectroscopy tools
Impurities	• Impurities	• Impurities from API, solvent and polymers	• Chemical stability of API, polymer and solvent during SDD process	• HPLC
Chemical stability	• Degradant level upon release and storage	• API physicochemical properties • Polymer, excipient, and solvent choice • Reactive impurities in polymers, excipient, and solvent	• Spray solution temp and hold time • Spray drying and secondary temperature and duration • Exposure to humidity and temperature	• HPLC
Water content	• Potency • Physical and chemical stability	• SDD, API and polymer hygroscopicity	• Secondary drying parameters • Exposure to humidity	• Karl Fischer
Residual solvent	• Residual solvent • Physical & chemical stability	• Solvent choice	• Droplet size and drying rate • Secondary drying parameters	• GC-MS • QNMR
Particle size	• In vivo performance • Downstream processing and final product quality attributes (eg, dissolution, assay, content uniformity, tablet hardness)	• Polymer chemistry and MW • Spray solution solids loading	• Droplet size • Particle attrition during agitated secondary drying	• Malvern • G3-Morphologi • Dissolution tests
Bulk density and morphology	• Downstream processing and final product quality attributes (eg, assay, content uniformity, tablet hardness)	• Polymer chemistry and MW • Spray solution solids loading	• Droplet size and drying rate	• Scotts volumeter • G3-Morphologi • SEM

30.3.1 API properties

The SDD performance attributes are highly dependent on the physicochemical and biopharmaceutical properties of the API. As exemplified next, a fundamental understanding of the impact of API properties on SDD performance can guide development of a robust SDD formulation. API properties that can impact SDD performance include: melting point (T_m); glass-transition temperature (T_g); Kauzmann temperature (T_k), the temperature at which the difference in entropies becomes zero; crystalline form (including counterions for salts); acid dissociation constant at logarithmic scale (pK_a); miscibility in polymers; hydrogen bond donors and acceptors; lipophilicity (eg, log P); epithelial membrane permeability; solubility in aqueous biorelevant media and spray solvents; and chemical stability.[29–34] Trasi et al. and Van Eerdenbrugh et al. recently demonstrated the importance of API molecular properties on ASD physical stability. The authors found that the crystal growth rates of compounds could vary by about five orders of magnitude between the fastest growing compounds and the slowest growing compounds.[33,35,36] The authors found that physicochemical properties of the compounds that had fast crystal growth rates had lower molecular weights, higher T_m values, lower T_g values, fewer rotatable bonds, lower melt entropies, lower melt viscosities, and higher crystal densities.

T_g is one of the important API characteristics that is used to predict the physical stability of the SDD. T_g is the temperature at which the amorphous solid converts from a glassy state (solid) to super-cooled liquid state, as shown in Fig. 30.3.[37] Structural factors, such as molecular size and shape, and the extent, strength, and direction of hydrogen bonding can affect T_g. A higher T_g value and a lower T_m/T_g ratio generally indicates that amorphous API will have a lower propensity to crystallize.[33] Generally, APIs with T_g values much higher than the storage temperature (T_s) will have a low tendency to crystallize, and SDDs made with such APIs tend to be physically stable.[38] In addition to the driving force for crystallization in the solid phase, the API lipophilicity (indicated by log P) is an important factor because highly lipophilic compounds often form SDDs that wet poorly and, consequently, dissolve poorly.

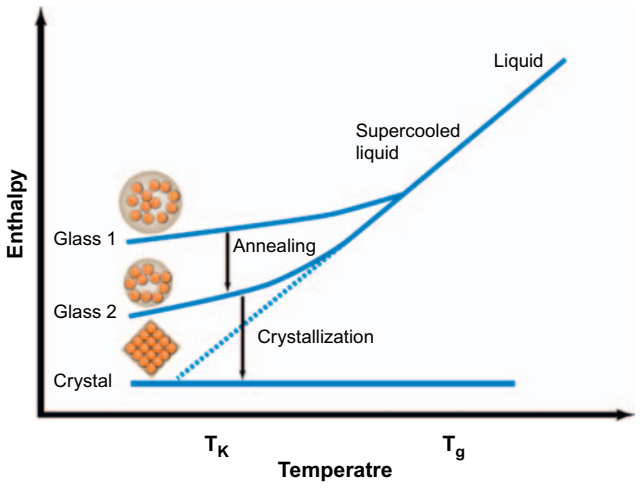


FIGURE 30.3 Temperature versus enthalpy (E). Depending on thermal history, glasses can form with slightly different energies, resulting in variable T_g values.

An API's tendency to crystallize and its lipophilicity can be used to guide the selection of the polymer and excipients and determine the maximum drug loading for an SDD.[32] In general, compounds that have higher tendency to crystallize require a higher ratio of polymer-to-drug to form SDDs that resist crystallization. Compounds with high log P values may also require a larger ratio of polymer to drug and excipients to reduce drug–drug interactions, improve wetting, and achieve the desired dissolution performance. Wetting and, hence, dissolution of SDDs can also be improved by incorporating surfactant in the SDD or in the dosage form.[39]

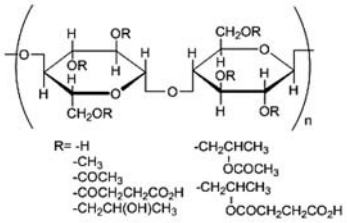
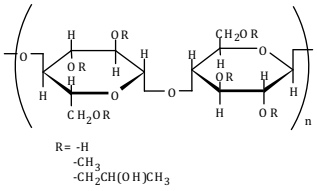
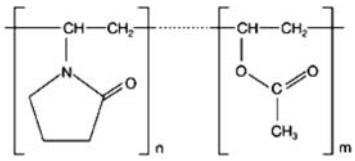
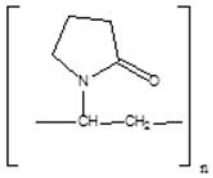
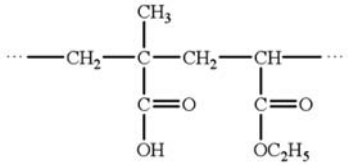
API is generally isolated in a crystalline form even if it is later converted to an amorphous dispersion for bioavailability purposes because the crystalline drug form has many processing advantages. Different drug forms (eg, salts, free forms, solvates) and polymorphs can have different solubilities in spray solvents, which can affect processing ease. Therefore, it is important to select a suitable crystalline drug form for SDD development and to evaluate the solubility of the selected form in the desired spray-drying solvent or solvent mixture. The API's tendency to form a solvate with the spray solvent should also be examined because these solvates will have lower solubility than the neat form. When dealing with acidic or basic drug compounds, acid–base equilibria can take place in the amorphous state, leading to unanticipated changes in various solid-state properties. In cases where a salt of the crystalline form is isolated and used to form an SDD, special attention should be paid to the choice of counterion because it can affect properties of amorphous materials. Towler et al. reported that use of counterions with high electrophilicity indices resulted in stronger intermolecular interactions, producing dispersions with higher T_g values.[40] Towler et al. also found that T_g is affected by the pK_a of the counterion, with higher T_g values observed for counterions with lower pK_a values.[40] Sonje et al. evaluated the effect of monovalent sodium and bivalent calcium and magnesium counterions on the properties of

amorphous salts of atorvastatin prepared by spray-drying.[30] They found that the choice of counterion affected the T_g, fragility (ie, temperature dependence of molecular mobility), crystallization tendency, and chemical stability of the drug.[30]

30.3.2 Polymer choice

Other than the physical and chemical properties of the API, the polymer has the most significant impact on SDD's ability to achieve and maintain supersaturation, physical and chemical stability, and powder properties.[41,42] SDDs employ polymers as a matrix material to achieve desired performance criteria (Fig. 30.2) that typically cannot be achieved with pure amorphous API. Key polymer properties that influence choice of dispersion polymer include: (1) miscibility with the drug over the desired range of compositions; (2) functional groups (eg, acidic, basic, and hydrogen bond donors/acceptors that can result in drug–polymer interactions); (3) T_g (higher is typically better); (4) low hygroscopicity; (5) aqueous solubility; (6) precipitation inhibition characteristics; and (7) solubility in organic solvents suitable for spray-drying. Table 30.2 lists common polymers used for SDDs.

TABLE 30.2 Polymers Commonly Used in Spray Dried Dispersions (SDDs)

Polymer (grade)[a]	Structure	T_g (°C)		Water content at 25°C/75% RH (wt%)	Typical spray solvents[c]	Ionizable groups
		<5% RH[b]	At 75% RH			
HPMCAS (M)		125	73	6	Acetone, methanol, THF, DCM	Acidic
HPMC (E3)		142	43	11	Methanol/water	None
PVP-VA (VA-64)		106	17	21	Acetone, methanol	None
PVP (K15)		116	−14	35	Methanol	None
PMA (Eudragit L100)		191	97	12	Acetone, methanol, DCM	Acidic

[a]*HPMCAS (M) = the M grade of hydroxypropyl methylcellulose acetate succinate, HPMC (E3) = the E3 grade of hydroxypropyl methylcellulose; PVP-VA (VA-64) = the VA-64 grade of polyvinylpyrrolidone vinyl acetate, PVP (K15) = the K15 grade of polyvinylpyrrolidone, PMA (Eudragit L100) = polymethacrylate marketed as Eudragit L100.*
[b]*RH—relative humidity.*
[c]*THF—tetrahydrofuran; DCM—dichloromethane.*

Polymers stabilize SDDs by reducing the molecular mobility of the drug by forming intermolecular interactions between drug and polymer, and by reducing the chemical potential of the drug.[43] An amorphous drug is typically most stable when drug and polymer are mixed homogeneously at the molecular level.[44,45] An advantage of spray-drying, as opposed to HME for example, is that rapid drying kinetics trap drug and excipients in a well-mixed state in the low-mobility polymer matrix. Therefore, in an SDD, the drug does not have to be miscible in the polymer to the level at which it is loaded. Nevertheless, the miscibility of drug in the polymer matrix can play a role in the physical stability of the dispersion with respect to phase separation, including crystallization. The strong interactions between an API and a polymer via ionic interactions, hydrogen bonding, halogen bonding, van der Waals forces (eg, dipole–dipole, dipole-induced dipole, London dispersion forces), and hydrophobic interactions are expected to facilitate miscibility of the drug in the polymer and may increase physical stability.[46–48] Thus, a molecular-level understanding of the interactions between the drug and polymer can aid in the selection of an appropriate dispersion polymer and drug loading for the SDD.

One of the key mechanisms by which polymers stabilize amorphous SDDs is by reducing drug mobility. Polymers with high T_g values are preferred to limit drug mobility and the resulting potential for drug-excipient phase separation. Typically, a molecular dispersion will have a T_g that is between that of the drug and polymer. The T_g of the drug–polymer mixture can be estimated using several approximations, including the Fox equation (Eq. (30.2))[49] and the Gordon-Taylor equation (Eq. (30.3)).[50] In these equations, T_{g1} and T_{g2} and w_1 and w_2 are the T_g values and weight fractions of drug and polymer, respectively, and K_{GT} is a constant that indicates interaction between two components. Eq. (30.4) shows the calculation of K_{GT}, with ρ_1 and ρ_2 being true densities (g/cm^3).

$$\frac{1}{T_g} = \frac{w_1}{T_{g1}} + \frac{w_2}{T_{g2}} \tag{30.2}$$

$$T_g = \frac{w_1 T_{g1} + K_{GT} w_2 T_{g2}}{w_1 + K_{GT} w_2} \tag{30.3}$$

$$K_{GT} = \frac{\rho_1 T_{g1}}{\rho_2 T_{g2}} \tag{30.4}$$

In many cases, the theoretically predicted T_g values for the drug–polymer system deviate from the experimental value. The reason for this deviation is nonideality in mixing and/or interaction of the drug with polymer, which could be because drug–drug and polymer–polymer interactions are different than the drug–polymer interaction. A strong or equal interaction between drug and polymer due to ionic interactions, hydrogen bonding, halogen bonding, van der Waals forces (eg, dipole–dipole, dipole-induced dipole, London dispersion forces), and hydrophobic interactions will lead to a more stable SDD.

Besides physical stability, polymers have a significant impact on the ability of the SDD to achieve and sustain supersaturation in solution.[51,52] The general concept behind almost all solubilization technologies is termed the "spring-and-parachute" effect, where the "spring" portion relates to the ability of the SDD to achieve supersaturation and the "parachute" portion relates to the ability of the SDD to maintain supersaturation long enough for drug absorption to take place.[53] Most of the polymers used for SDDs are hydrophilic or at least amphiphilic (Table 30.1). Charged or ionizable polymers include substituted acrylates (eg, polymethylacrylate (PMA) and hydroxypropylmethylcellulose acetate succinate (HPMCAS)). Uncharged polymers include hydroxypropylmethylcellulose (HPMC), polyvinylpyrrolidone/vinyl acetate (PVP-VA), and polyvinylpyrrolidone (PVP). Neutral water-soluble polymers can be useful for applications where rapid dissolution of the drug in the stomach is needed and where sustainment of supersaturated drug is not made difficult by a drug's strong tendency to crystallize. The rapid dissolution can be advantageous to maximize the time of exposure of high concentrations of dissolved drug in the small intestine.

Enteric polymers, such as HPMCAS, can be useful for applications where rapid gastric dissolution in the gastric is not needed or desired. For example, the high solubility of some weakly basic drugs in the low-pH gastric environment can result in high supersaturated drug levels, especially upon transit and subsequent neutralization of the drug in the higher pH of the small intestine, which can lead to rapid crystallization. In these instances, it can be advantageous to use an enteric polymer in the SDD to retard dissolution of the drug until the material reaches the small intestine, where the supersaturation will be more moderate and rapid crystallization of the drug to a lower-energy form will be less likely.[32,54]

On the other hand, depending on the SDD particle-size distribution and on the specific interactions between the drug and polymer in a hydrated environment, drug release can be rapid without concomitant dissolution of the enteric polymer. In this scenario, use of the enteric polymer can be a disadvantage if the supersaturated drug precipitates before the enteric polymer can dissolve in the higher pH of the small intestine and interact favorably with the drug to inhibit crystallization.

For acidic APIs that may have low solubility in the gastric environment and higher solubility in the

higher-pH intestinal environment, an enteric polymer is often not necessary to prevent premature release and supersaturation. In these cases, use of a neutral water-soluble polymer can sometimes result in more rapid release when a drug travels to the small intestine. Therefore, the choice of optimal polymer for the SDD is often dictated by a number of drug properties, including solubility, ionizability, lipophilicity, and propensity to crystallize, as well as the required SDD drug loading.

The impact of polymer selection on the in vivo performance of SDDs has been reported in the literature.[13,42,55] Qian et al. described the impact for SDDs made with BMS-A and PVP-VA or HPMCAS.[42] In this case, BMS-A was miscible with both PVP-VA and HPMCAS, with the former better solubilizing the drug. The in vitro dissolution performance of both SDDs was similar, as shown in Fig. 30.4, but in in vivo testing, the BMS-A/HPMCAS SDD had a higher maximum drug concentration (C_{max}) and area under the time-concentration curve (AUC) than the BMS-A/PVP-VA SDD, as shown in Fig. 30.5. The authors found that the fast dissolution of the PVP-VA resulted in elevated drug loading in partially undissolved particles and facilitated drug recrystallization before complete release. In contrast, the hydrophobicity and slower dissolution of the HPMCAS prevented BMS-A recrystallization within the HPMCAS matrix, effectively prolonging BMS-A supersaturation.

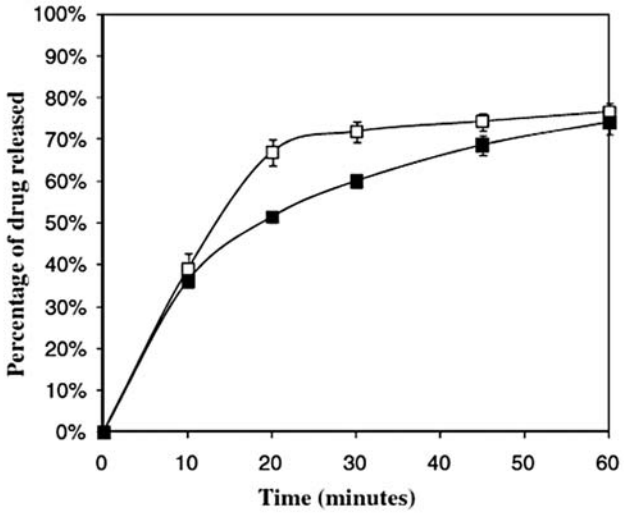


FIGURE 30.4 In vitro release kinetics of BMS-A amorphous formulations. USP II sink condition dissolution of PVP-VA (□) and HPMCAS (■) solid dispersion. Dissolution medium: 1000 mL of pH 4.5 acetate buffer with 1.5% Brij 35. Paddle speed 50 rpm. *Source: From Qian F, Wang J, Hartley R, Tao J, Haddadin R, Mathias N, et al. Solution behavior of PVP-VA and HPMC-AS-based amorphous solid dispersions and their bioavailability implications. Pharm Res 2012;29:2765–76. Published with permission of Springer.*

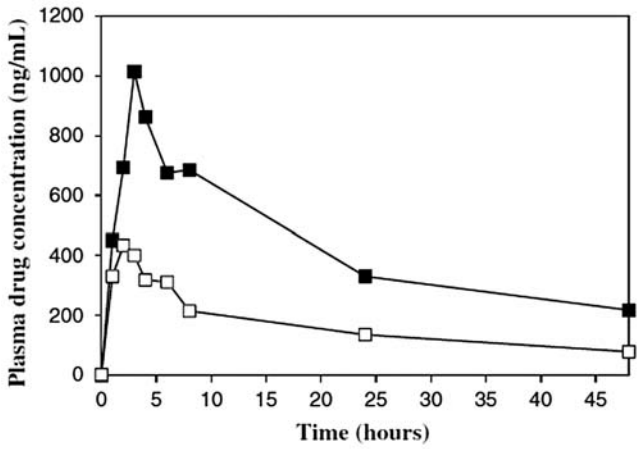


FIGURE 30.5 In vivo pharmacokinetic performance of BMA-A amorphous solid dispersions. BMS-A/PVP-VA solid dispersion (□) and BMS-A-HPMC-AS solid dispersion (■) in a dog PK study ($n = 4$, 75 mg/dog). *Source: From Qian F, Wang J, Hartley R, Tao J, Haddadin R, Mathias N, et al. Solution behavior of PVP-VA and HPMC-AS-based amorphous solid dispersions and their bioavailability implications. Pharm Res 2012;29:2765–76. Published with permission of Springer.*

The enteric polymers tend to be some of the more amphiphilic polymers used in SDDs and therefore form drug–polymer colloids more readily. Such colloids provide a high-activity, high-surface area source of drug that can readily replenish a dissolved drug as it is absorbed.[32] In addition, colloids may act to shuttle the drug across the unstirred boundary layer. On the other hand, acidic polymers can be incompatible with drugs that have acid-sensitive chemical stability liabilities. For these drugs, it can be advantageous to use nonacidic polymers to avoid the potential for chemical instability. In some cases, modestly reducing the number of acid groups on the polymer—for example, changing from the (L) grade to the (M) or (H) grade of HPMCAS—can have a significant positive effect on chemical stability.[56]

30.3.3 Additional excipients

Drug and polymer are typically the two primary components of an SDD, but sometimes additional excipients are included in the SDD formulation. These excipients might include antioxidants to improve the oxidative stability of the drug in the amorphous state; pH modifiers to improve stability or mitigate a pH effect[57,58]; superdisintegrants to improve disintegration/dissolution; glidants to improve bulk material properties[59,60]; complexing agents (eg, cyclodextrins) to bind and solubilize drug[61]; or surfactants to improve wettability and maintain supersaturation.[62,63]

The addition of other excipients, such as surfactants or lipids, to the SDD can be advantageous in specific

circumstances. For example, addition of low-molecular-weight surface-active agents such as sodium lauryl sulfate (SLS) or Vitamin E d-alpha tocopheryl polyethylene glycol 1000 succinate (TPGS) can enhance dissolution rates for some SDDs, particularly those containing highly lipophilic drugs, and may also promote the formation of mixed micelles with bile salts in the small intestine to increase the micellization of drug.[64–66] In addition, some low-molecular-weight surface-active agents, such as Vitamin E TPGS and polyethylene oxide oligomers (PEO) have been reported to interact with the gut epithelium to enhance absorption.[67] The addition of excipients can change solution-state interactions and can affect the spray-drying process, which in turn can influence SDD particle properties (eg, shape, morphology, physical structure, surface chemistry) and performance. Therefore, it is important to understand the interactions at the molecular level where possible, and their impact on SDD processability, performance, and stability. SDD process involves solubilization of all formulation components in solvent that is removed during the manufacturing process. Solvent is not considered an excipient, but its choice and residual level can have a plasticizing effect that potentially influences the SDD's physical and chemical stability.

30.3.4 Drug loading

The target drug exposure is a critical factor, as dose and bioavailability are interrelated, and therefore impact the drug loading requirements for the SDD. The TPP is typically based on the size and a maximum number of dosage units to deliver the target dose.[27,28] Generally, a higher drug loading (drug to polymer ratio) in the SDD is desirable to minimize the final product size and units. The maximum drug loading in the SDD will depend on the physical and chemical stability, dissolution performance, and powder properties of the SDD as a function of drug loading. Typical achievable active loadings are 25–50%A, but in the authors' experience, loadings from 1%A to 90%A have been used.[68] The maximum achievable loading is often limited for drugs with high T_m and low log P values that have a strong tendency to crystallize from the amorphous state, as shown in the map in Fig. 30.6.[32] Dilution of these drugs in the excipient matrix can help promote physical stability during spray-drying, storage of the SDD in the solid state, and help limit precipitation when the SDD is introduced into the use environment. The maximum drug loading for a particular drug depends not only on the drug properties, but also on the characteristics of the polymer. Polymers in which the drug is more miscible or that offer a lower mobility environment can help stabilize the drug against crystallization or amorphous phase separation. In many cases, the drug has a lower T_g than the polymer. In these cases, the T_g of the SDD typically decreases as drug loading is increased, which can increase the risk of phase separation due to the potential for increased drug mobility at typical storage conditions (ie, temperature and relative humidity, RH).

High drug loadings in the SDD can also result in poor dissolution properties, especially for highly lipophilic drugs that may wet poorly in the aqueous milieu of the GI tract. Often, the dissolution mechanism of such SDDs is by slow erosion. In such cases, a lower drug loading and a water-soluble polymer can be used to enhance the dissolution rate through faster erosion or by changing the dissolution mechanism from erosion to disintegration. These mechanisms are described in more detail later.

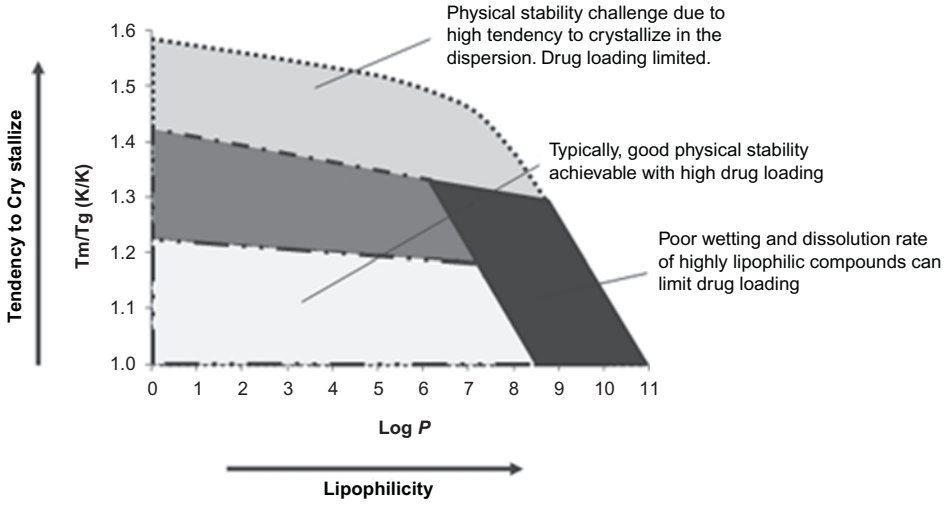


FIGURE 30.6 Physical stability map for SDDs, showing crystallization tendency as a function of lipophilicity. *Source: Modified from Vodak DT, Morgen M. Design and Development of HPMCAS-Based Spray-Dried Dispersions. In: Shah N, Sandhu H, Choi DS, Chokshi H, Malick W, editors. Amorphous Solid Dispersions. New York: Springer; 2014. p. 303–22.*

30.3.5 Spray solvent

Since the choice of solvent has a large impact on process and product quality and performance, it is important to identify a solvent that: (1) can dissolve the polymer (which is generally hydrophilic), the drug (which is generally hydrophobic), and any additional excipients at sufficient concentration; (2) has high volatility that enables rapid drying of spray-droplets; (3) is chemically compatible with the formulation components and spray-drying equipment; (4) has acceptable rheological properties; (5) is chemically stable; (6) is nontoxic[69]; (7) has acceptably low levels and types of impurities; (8) is noncombustible under spray-drying conditions; and (9) is eco-friendly. The solvent-dependent parameters, such as evaporation rate, diffusion coefficient, API/polymer solubility, API/polymer/solvent interactions, rheological properties, and chemical compatibility can have a significant impact on SDD characteristics such as particle morphology, particle density, particle size, API-polymer miscibility, yield, dissolution, residual solvent, and physical and chemical stability.[70] Table 30.3 provides some of the relevant physical properties for common spray solvents.[29] In many cases, a single solvent with the desired characteristics cannot be identified, so a mixture of solvents may be used to achieve the desired solution characteristics.

The solubilization capacity of the solvent is highly dependent on its dielectric constant, dipole moment, total solubilization parameter, and hydrogen bonding characteristics.[29,71] The development and optimization of the spray-drying process requires screening of the API solution solubility as a function of temperature and solvent composition using manual or high-throughput screening systems. Due to the structural requirements needed to address new drug targets, many of the drugs in current development programs are poorly soluble not only in aqueous media, but also in typical spray-drying solvents. Rigorous solvent screening protocols involving a combination of experimental and theoretical methods[72] are required to identify optimal spray-drying solvents in which the drug has sufficient solubility for acceptable throughput, yield, and particle properties.

Solvents can influence the conformational properties of polymer and, thus, the properties and physical stability of the SDD.[29,70,73,74] A polymer in a "good" solvent will typically be better solvated and "extended" than the same polymer in a "bad" solvent, where it might assume a more "compact" structure. Al-Obaidi found that spray-drying griseofulvin, poly[*N*-(2-hydroxypropyl) methacrylate] (PHPMA) and PVP from an acetone/methanol solvent system produced SDDs with dramatically different morphology, stability, and dissolution properties compared with a similar SDD prepared using acetone/water as the solvent system.[73] The authors attributed the differences in SDD properties to the conformational variations of the polymers in solution before spray-drying. Al-Obaidi et al. and Paudel et al. showed that choice of a poor spray-drying solvent can impact SDD phase behavior and, thus, the physical stability and in vitro/in vivo performance of the SDDs.[70,73]

In a case study involving a BMS-A/PVP (K30) SDD, the authors (Qian F, AAPS Webinar, March 2011) showed that BMS-A and PVP (K30) were miscible based on the negative Flory–Huggins interaction

TABLE 30.3 Relevant Properties of Common Spray-Drying Solvents

Solvent	Boiling point (°C)	Density (g/mL)	Solubility in water (g/100 g)	Relative polarity	ICH limits[a] (ppm)	ΔH_v[b] (kJ/mol)	Vapor pressure[c] (hPa)	Viscosity[d] (mPas)	Surface tension[d] (mJ/m^2)
Acetone	56.3	0.786	M[e]	0.355	Class 3	29.1	240	0.3	22.68
DCM	39.8	1.326	1.32	0.309	600	28.0	475	0.42	27.20
Ethanol	78.3	0.789	M	0.654	Class 3	38.7	59	1.08	22.00
Ethyl acetate	77.1	0.894	8.7	0.228	Class 3	31.9	97	0.43	23.20
Isopropanol	82.3	0.785	M	0.546	Class 3	45.7	44	2.07	18.30
Methanol	64.7	0.791	M	0.762	3000	35.3	128	0.54	22.10
THF	66.0	0.886	30	0.207	720	26.9	200	0.46	26.40
Water	100	0.998	M	1	NA[f]	40.7	17.5	0.89	71.90

[a] *ICH = International Conference on Harmonisation.*
[b] ΔH_v *= enthalpy of vaporization.*
[c] *At 20°C.*
[d] *At 25°C.*
[e] *M—miscible.*
[f] *NA—not applicable.*

parameter. Further, infrared (IR) and solid-state nuclear magnetic resonance (ssNMR) spectroscopy showed nonlinear interaction between BMS-A and PVP (K30), indicating drug–polymer specific interaction. A 50/50 (w/w) BMS-A/PVP (K30) SDD was made using either an acetone/water solvent system or pure methanol as solvent. The SDD sprayed from the acetone/water solvent system showed two T_g events (at 78°C and 128°C), whereas the SDD sprayed from methanol had a single T_g of 90°C. Further, the SDD prepared with acetone/water showed the presence of crystalline API peaks when stored for 2 weeks at 25°C/60% RH in open containers, whereas no signs of crystallization were seen when the SDD prepared from methanol was stored under similar conditions.

Solvent properties such as boiling point, enthalpy of vaporization, and vapor pressure influence the energy required and, thus, the process parameters needed to remove the solvent during spray-drying. Rheological properties, such as viscosity and surface tension, affect the atomization of the spray solution. The composition and the viscosity of the spray solution significantly affect droplet formation, drying, and the particle properties of the dry SDD.[75]

The solvent's thermal properties largely influence the residual solvent level. Solvent levels should be controlled per ICH guidelines because an excess of residual solvent can be toxic.[69] In addition, the residual solvent can plasticize the SDD, reducing its T_g and potentially causing physical instability.[76] Residual solvent levels are usually reduced to well below the recommended ICH guidelines. The vast majority of the solvent is removed during the spray-drying process. A secondary drying process (the fourth step in the spray-drying process) is then used to remove residual solvent to the desired level. The solvents used in the spray-drying process often contain impurities that can be more toxic than the solvents themselves. For example, acetone is often tainted with mesityl oxide (MO) and diacetone alcohol (DAA), and methanol can contain trace levels of benzene. During the spray-drying process, the higher boiling points of these impurities can result in their concentration in the SDD. Therefore, it is usual to assess levels of potential impurities, especially if they are toxic. In one case, Quirk et al. developed a sensitive and validated headspace gas chromatography (GC) method to determine residual levels of acetone and acetone-related impurities (ie, MO and DAA) in SDDs.[77] Yue et al. recently proposed a control strategy for benzene impurity in HPMCAS-based SDDs. To eliminate risk of residual benzene reaching patients, the authors developed a science- and risk-based approach to establish upstream control of benzene in solvents and in-process control of residual solvents in polymer-stabilized SDDs.[78]

30.4 SDD PROCESS CONSIDERATIONS: MANUFACTURING AND SCALE-UP

The spray-drying process involves atomizing and drying solution of drug, polymer, and potentially other excipients in a volatile spray solvent. As shown in Fig. 30.7, the spray-drying process typically involves four steps: (1) preparation of spray solution, (2) atomization and drying of spray solution, (3) collection of

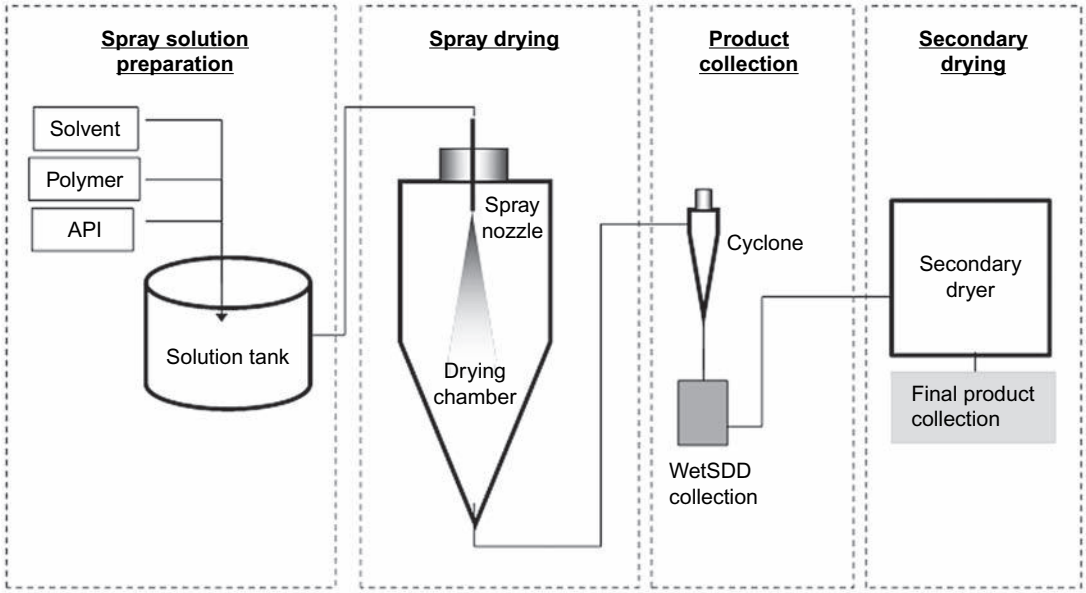


FIGURE 30.7 Spray-drying process overview.

the wet (ie, solvent-laden) SDD, and (4) secondary drying of the wet SDD to remove residual solvent.

Spray dryers are available in a variety of scales: lab units that can process milligram quantities, pilot units that can process gram to kilogram quantities, and commercial units that can process metric tons. Although the fundamental principles for the spray-drying process remain the same, the large units are more efficient and the powder properties of the materials from small units can be significantly different from those produced at larger scale.

30.4.1 Spray-solution preparation and considerations

Spray-solution preparation is the first step of the spray-drying process in which all the components—drug, polymer, and any additional excipients—are dissolved in solution prior to atomization. A number of considerations during this step can impact SDD quality attributes and processability. These include: (1) the solvent choice; (2) the solution composition and solids content (ie, API, polymer, and excipient concentration); (3) solution rheological properties (eg, viscosity, surface tension, and density); (4) chemical stability (of spray solution, including solvent); (5) polymer conformation; (6) solution mixing (especially for processes in which the spray solvent is heated to facilitate dissolution of components and increase organic solubility); and (7) spray solution stability. Selection of the spray solvent is very important for generation of SDDs with desired stability and material properties. The key considerations during selection of the spray solvent were covered in the previous section.

The achievable spray solution concentration is largely dependent on the solubility of API in spray solvent because the solubility of common SDD polymers is sufficiently high in common spray solvents. Typically, spray solution solids concentration ranges from 2–20% w/v. The spray solution concentration has a direct impact on its rheological properties and thus, the droplet formation and drying. Increase in spray concentration generally results in bigger and denser particles. This is due to higher spray solution viscosity, less liquid to evaporate, and lower diffusion of solutes within the droplet. In contrast, dilute spray solutions typically results in smaller and lower density particles.[70] However, an increase in the spray concentration can result in phase separation, while lower spray concentration can increase hygroscopicity (and hence physical instability) due to higher effective surface area of the SDD.[70,73] Therefore, when developing a spray-drying process, the impact of spray solution concentration of droplet size, drying, and SDD physical stability and performance should be evaluated.

The spray-drying process can take several hours to days, depending on the manufacturing scale. Therefore, it is important to establish the stability of spray-solution for an appropriate duration (generally several days) and temperature range. The spray solution hold-time should cover manufacturing time and additional time to cover unforeseen issues that might arise during manufacturing. Chemical stability of the API in the presence of other spray solution components (polymer, excipients, and solvent) is generally of highest concern. The likelihood of API degradation due to reactive impurities is high because all components are in solution. In addition to chemical stability, physical instability due to API form change (eg, solvate formation), polymer aggregation should be monitored. The spray solution temperature should be properly controlled or their impact understood because small changes in temperature can impact stability and rheological properties (and thus, SDD powder properties). Physical stability and degradation kinetics of spray solution at various temperatures should be evaluated during lab stability studies to enable hold-time for spray solution.

30.4.2 Warm and hot spray drying processes

Drugs with high T_m values can have limited solubility in organic spray solvents, resulting in a spray-drying process with unacceptably low throughput, as well as the production of particles with poor properties (eg, small size). To efficiently process such high-T_m compounds, the temperature of the spray solution can be increased to increase the solubility of active in the spray solvent. In some cases, only a modest increase in the solvent temperature—such that the solvent is still below its boiling point—is required. This is known as a "warm process." In other cases, especially where a large enhancement in solubility is required, the spray-solution temperature is increased to well above the ambient-pressure boiling point of the solvent.[79] In this high-temperature spray-drying "hot process" the drug is dissolved immediately before it is introduced into the spray-dryer, as shown in Fig. 30.8. The drug suspension is constantly stirred to ensure homogeneity and an in-line heater exchanger is used to heat the suspension above the solvent boiling temperature. The pressure of the solution is controlled to prevent the solution from boiling before it exits the heat exchanger and is fed through the spray nozzle. Typically, the temperature is chosen so that the concentration of drug is 30–50% of the solubility of the drug in the spray solvent at that temperature to ensure production of homogenous particles. A flash nozzle is typically used to achieve solution atomization.[80,81] The drying resulting from the hot process is generally faster than for the standard

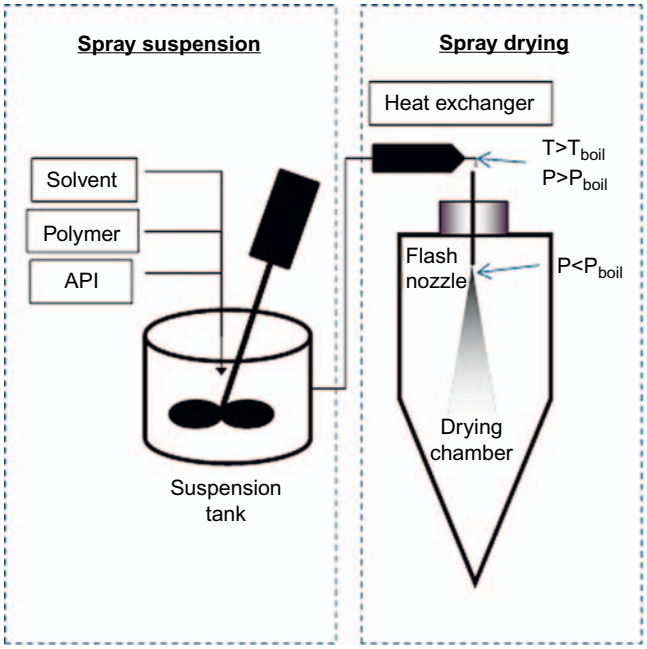


FIGURE 30.8 Schematic of SDD manufacture by hot process (*T*, temperature; *P*, pressure).

spray-drying process, resulting in shorter solidification times for the particles; which can be advantageous for making homogeneous dispersions.

30.4.3 Atomization and drying of spray solutions

The spray-drying process starts with atomization of the spray solution into droplets, which dry when they contact heated gas in the drying chamber. Atomization of the spray solution depends on (1) spray-solution flow rate and pressure, (2) nozzle geometry, and (3) spray-solution viscosity and surface tension. Drying of the atomized droplets depends on (1) droplet size, (2) inlet temperature, (3) outlet temperature, (4) relative saturation at the spray-dryer outlet, (5) and the specific drying ratio (ie, the mass ratio of liquid to drying-gas flow rate). These steps and process variables are described next.

30.4.3.1 Atomization

The purpose of atomization is to convert the spray solution into fine droplets. The resulting increase in the surface area improves efficiency for heat and mass transfer, promoting rapid drying of droplets in the spray-drying chamber. Due to rapid atomization, the drying material never reaches the inlet temperature of drying gas, which makes spray drying a method of choice for heat-sensitive APIs. SDD composition and process can be tuned to allow control of particle size and morphology.[75,82]

Atomization of the spray solution is achieved by using scale-appropriate spray nozzles. There are several types of spray nozzles including: (1) rotary atomizers, (2) two-fluid or pneumatic nozzles, (3) pressure nozzles, (4) ultrasonic nozzles, and (5) flash atomizers.[83,84] Rotary atomizers (disk or wheel type) achieve atomization by discharging liquid from the edge of a wheel rotating at high speed (ie, 10,000–50,000 rpm). These nozzles, which have a spray angle of about 180 degree, are generally used in wide-diameter spray-dryers at large scale. Rotary nozzles typically produce droplets in the range of 20–200 μm, with the droplet size being inversely proportional to the disk or wheel speed. The high energy provided by the rotary nozzles makes them ideal for atomizing slurries, suspensions, and high-viscosity solutions. Two-fluid or pneumatic nozzles use pressurized gas to break the liquid feed into small droplets. These nozzles typically produce droplets in the range of 5–75 μm. These nozzles provide very good control over the droplet size because the spray-solution feed rate and gas flow rate can both be controlled independently. One disadvantage of these nozzles is that the use of pressurized gas for atomization can limit their use at industrial scale.

Pressure nozzles generate droplets by pressurizing the spray solution and forcing it through a nozzle orifice. These nozzles generally include a chamber or swirl insert that imparts swirling motion to the liquid inside the chamber. The swirling motion pushes the liquid to the wall in a hollow sheet, leading to the desired hollow cone spray pattern when the spray solvent is ejected. These nozzles typically produce droplets in the range of 30–200 μm. The droplet size is inversely proportional to the atomization pressure and directly proportional to the spray-solution feed flow rate. However, these two parameters are interdependent because an increase in atomization pressure also increases spray-solution feed-flow rate. The drawback of this interdependency can be overcome with the multitude of commercially available nozzle geometries.

Ultrasonic nozzles produce droplets in a fine mist spray using high-frequency sound vibrations (ie, 20–120 kHz). The liquid spreads over a vibrating small horn after exiting the nozzle orifice, generating a uniform, narrow droplet-size distribution ranging from 20 to 100 μm in a low-velocity spray. However, the limited throughput of these nozzles limits their use at commercial scale. Flash atomization occurs when a solution that is heated above the boiling point at ambient temperature reaches the equilibrium vapor pressure. Boiling and nucleation can start and the energy from this process drives droplet formation.[80] Flash atomization is not often used for SDD atomization, although it is preferred while using the hot process.

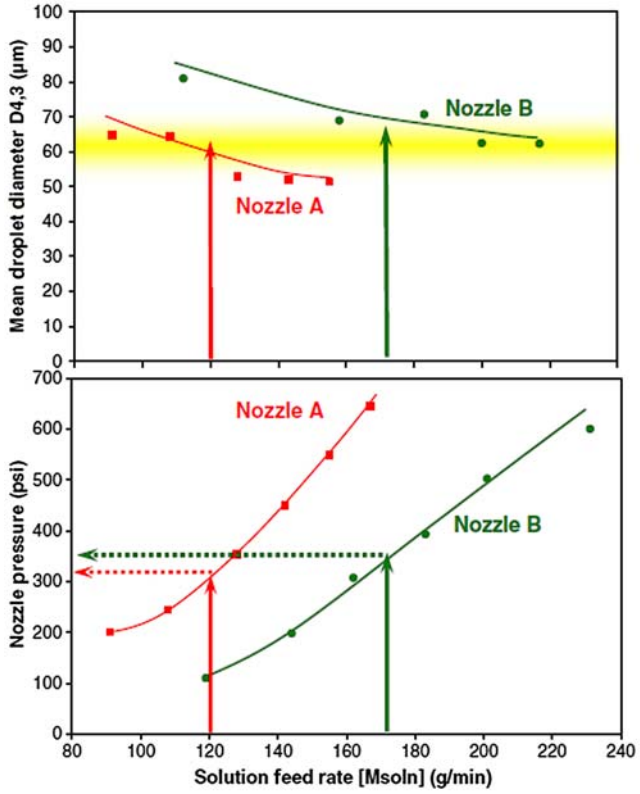


FIGURE 30.9 Results of phase doppler particle analyzer (PDPA) testing, showing representative droplet diameters and nozzle pressures as a function of solution feed rate. *Source: From Dobry DE, Settell DM, Baumann JM, Ray RJ, Graham LJ, Beyerinck RA. A model-based methodology for spray-drying process development. J Pharm Innov 2009;4:133–42. Published with permission of Springer.*

The choice of atomizer is dictated by the spray-dryer scale, solution properties, throughput, and desired particle-size distribution. In SDD production, pressure nozzles are preferred for their (1) ease of use and cleaning, (2) robust operation, (3) ability to handle variety of feeds with broad rheological properties, (4) narrow droplet-size distribution, and (5) scalability. Pressure nozzles have been well studied and characterized.[85] The size of the dried particle is directly correlated with the size of the droplet and the solids content of the spray solution, as described by Eq. (30.5)[86]:

$$D_{\text{particle}} \cong D_{\text{droplet}} \times \sqrt[3]{x_{\text{solids}} \times \frac{\rho_{\text{droplet}}}{\rho_{\text{particle}}}} \qquad (30.5)$$

where D_{particle} is the diameter of the dried particle; D_{droplet} is the diameter of the droplet; x_{solids} is the mass fraction solids in the spray solution (ie, API, polymer, and excipient concentration); ρ_{particle} is the density of the dried particle; and ρ_{droplet} is the density of the spray solution. This relationship assumes a solid sphere morphology of a fixed density.

Droplet size can be measured experimentally with a phase Doppler particle analyzer (PDPA), using either the actual spray solution or a placebo solution with matching rheological properties. Fig. 30.9 shows representative droplet sizes produced using two pressure nozzles at various atomization liquid pressures and solution feed rates.[86] Such data can be used to select a nozzle and atomization parameters that yield the desired particle-size distributions.

30.4.3.2 *Drying*

Particle are formed during spray-drying by atomizing the feed solution to form droplets and subsequent drying of the droplets. The atomized droplets mix with the drying gas maintained at the appropriate temperature and humidity to dry the solvent from the droplets. To maximize the heat and mass transfer, the atomizer and drying-gas inlets are positioned at the top of the drying chamber and liquid spray-solution feed and drying-gas flow are cocurrent. As the droplets begin to dry, a film forms around the outer edges of the droplet in a process known as "skinning." As drying progresses, the remaining solvent is driven off, increasing the solid density and generating the SDD particle. Following drying of the solvent, a particulate aerosol forms. Depending on the gas flow, these particles may recirculate in the drying chamber for some time before exiting into a cyclone separator and/or bag filtration system for separation of the solid SDD particles from the drying gas. The drying gas that exits the spray-dryer must be filtered to avoid contamination.

On larger pilot-scale and commercial systems, the drying gas then continues through a condenser unit, which removes much of the solvent vapor. The drying gas is then reheated and continuously recycled for the spray-drying process. Process parameters that have a large influence on the properties of the spray-dried SDD particles include: (1) the solids content of the spray solution; (2) the drug–polymer ratio; (3) the properties of the drug and polymer; (4) drying rate; (5) the outlet temperature of the spray-dryer; and (6) the identity and relative saturation of the spray-drying solvent.

30.4.3.3 *Scale-up considerations*

Improper atomization and drying can lead to a number of issues during spray drying, some of which are mentioned next.

30.4.3.3.1 Atomization

Improper atomization can lead to undesired droplet sizes, SDD particles with undesirable properties, and dripping of spray solution. Improper atomization conditions can produce droplets that are too small or too large. Small droplets will result in SDD with a large fraction of fines (small particles less than $\sim 5\,\mu m$) and poor flow properties. Large droplets are difficult to dry because they possess enough momentum to escape the gas whirl and deposit on the equipment

walls. Insufficiently dried material can phase-separate and be physically instable. Another issue with improper atomization is filament formation, which can happen if the viscosity of the spray solution is too high or if atomization energy is too low.

30.4.3.3.2 Product accumulation

Product accumulation on the walls of the equipment (eg, spray-drying chamber, cyclone, collection bags, ducts, filters) is a common issue. Product buildup is related to nonoptimal spray-drying parameters and may result in wet SDD with a higher residual-solvent content and lower T_g. The problem can be mitigated by optimizing the process conditions—for example, by either reducing the spray-dryer outlet temperature or by reducing outlet relative saturation by reducing condenser temp or reducing solution feed rate.

30.4.3.3.3 Bearding

Buildup of the product on the nozzle tip and around the nozzle (also known as bearding) is another common issue. Bearding can interfere with spray formation and droplet size—and, thus, product quality—and can clog the nozzle A number of approaches to overcome bearding, including repositioning the nozzle, changing the atomization conditions, using an additional gas stream concentric to the nozzle tip, and using an antibearding cap.

30.4.3.3.4 Condensation

Relative saturation of the drying gas can be significant when operating in recycle mode, particularly if water is used as a cosolvent. Condensation in the spray-drying chamber should be avoided by selecting conditions that will not encroach on the dew point, particularly as the drying gas leaves the insulated spray-drying chamber. During its transit, the drying gas may cool down further from the target process conditions.

30.4.4 Secondary drying

After spray-drying, the solid particles in the drying chamber may contain significant residual solvent, The residual solvent content of this so-called "wet" SDD must be reduced to a concentration below the ICH guideline.[69] Removal of residual solvent is also desirable because the presence of this solvent can plasticize the SDD and lower its T_g, leading to phase separation of the API and polymer.

Residual solvent is removed using the secondary drying process. This commonly involves vacuum desiccation at small scale, and tray-drying or agitated vacuum drying at larger scales. Agitated vacuum drying can, at least in principle, cause SDD particle attrition, which can affect powder properties and subsequent downstream processing into solid dosage forms.[68] Therefore, it is useful to understand the secondary drying mechanism and kinetics, as well as how the drying process affects the SDD particle quality attributes.

Hsieh et al. recently elucidated the fundamental drying mechanisms, developed a mathematical mass-transfer model, and used this model to predict the fate of organic solvents and solvent-related impurities, such as benzene in SDD products.[87] The authors concluded that the kinetics of solvent removal during secondary drying follows a Fickian diffusion model described by:

$$\frac{M_t}{M_0} = \sum_{n=0}^{\infty} \frac{8}{(2n+1)^2\pi^2} \exp\{-(2n+1)^2 t\beta\} \qquad (30.6)$$

and

$$\beta = \frac{D\pi^2}{4l^2}$$

where M_t is the weight of solvent in the SDD product at drying time t, M_0 is the weight of solvent in the wet SDD before secondary drying, n is an integer, and β is the diffusion parameter that is described by the diffusion constant (D) and diffusion length (l). The ratio of M_t to M_0 represents the fraction of solvent remaining in the wet SDD. As shown in Fig. 30.10, at the beginning of the diffusion ($t = 0$), the ratio of M_t to M_0 equals 1, and when the solvent is completely removed

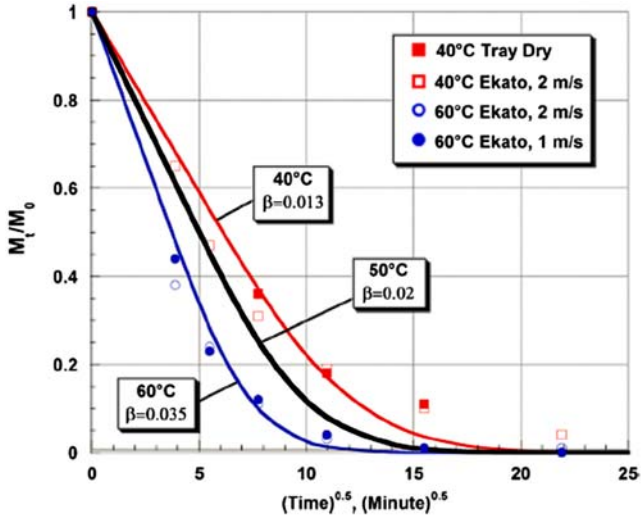


FIGURE 30.10 Representative results for scale-up of a secondary-drying process at various temperatures and drying scales for SDD particles. Conditions: SDD particles dried using a tray dryer and an Ekato agitated-vacuum dryer at 1 and 2 m/s. *Source: From Hsieh DS, Yue H, Nicholson SJ, Roberts D, Schild R, Gamble JF, et al. The secondary drying and the fate of organic solvents for spray dried dispersion drug product. Pharm Res 2015;32:1804–16. Published with permission of Springer.*

from the wet SDD, the ratio equals 0. For a given value of M_t/M_0, the secondary drying time will increase with diffusion length (l) and decrease with increase in the diffusion coefficient (β). Authors demonstrated that the diffusion coefficient can be sensitive to the secondary-drying temperature, depending on the relationship between T_g and the residual-solvent content of the SDD.

When a mixture of solvents (A and B) is present in a wet SDD, the relative diffusion rate of the two solvents can be described by Eq. (30.7). Where α_{AB} is the relative diffusion rate of Solvents A and B, β_A and β_B are the diffusion coefficients of Solvents A and B, D_A and D_B are the diffusion constants for Solvents A and B, and l^2_A and l^2_B are the diffusion lengths of Solvents A and B. Assuming that both solvents have the same diffusion length, the relative diffusion rate (α_{AB}) is simply the ratio of the diffusion coefficient of Solvent A (β_A) to the diffusion coefficient of Solvent B (β_B). This equation can also be applied to predict the fate of residual-solvent impurities in the primary spray solvent during the secondary drying process. Using this model, Hsieh et al. successfully predicted levels of residual benzene, an impurity in methanol, during the spray-drying process.[87]

$$\alpha_{AB} = \frac{\beta_A}{\beta_B} = \frac{D_A l_B^2}{D_B l_A^2} = \frac{D_A}{D_B} \tag{30.7}$$

30.4.5 SDD process and impact on bulk material properties

SDD material properties, such as SDD particle-size distribution, particle density, and morphology and derived properties such as flow, bulk density, and compressibility are dependent on SDD feed solution composition (API-polymer ratio, polymer chemistry, solvent, spray solution solid content) and SDD process parameters (droplet size, drying rate, and secondary drying).[88] As described previously, particle formation during spray-drying is achieved by atomizing the feed solution to form droplets, the size of which are dependent on the nozzle type, spray pressure, spray rate, and feed-solution viscosity, density and surface tension.[89] As the droplets begin to dry, a film is formed around the outer edges of the droplet in a process known as "skinning." As drying progresses, the remaining solvent is driven off, increasing the solid density and thereby creating void space within the SDD particle. The properties of API and excipients and the drying rate of the process determine the morphology of the SDD particles, which can range from hollow spheres to more dense, raisin-like structures.[75,90] Fig. 30.11 shows the tensile strength of particles as a function of compression strength for SDD dried at three sets of conditions. The scanning electron micrography (SEM) images show the effect on particle

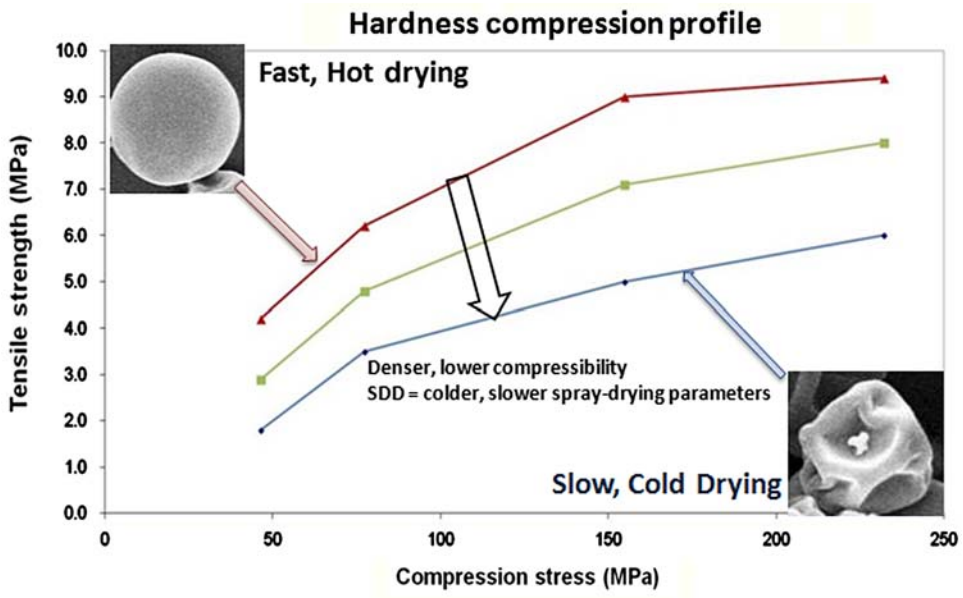


FIGURE 30.11 Effect of spray-drying conditions on SDD particles, showing tensile strength as a function of compression strength for three sets of conditions.

morphology for fast-hot drying conditions and slow-cold drying conditions.

Fig. 30.12 shows the effect of particle size on dissolution rate that can sometimes occur for SDDs prepared on different size spray-dryers (ie, laboratory, PSD-1, and PSD-2 scales). At the PSD-1 scale, SDD batches with two sets of solids contents are shown.

30.4.6 Scale-up considerations

When spray-drying processes are scaled up, the goal is often to maintain the SDD properties attained at small scale (eg, particle-size distribution, particle density and morphology, and the homogeneous amorphous state), while achieving the throughput needed for a later stage drug development program. Scale-up typically affects the atomization and droplet drying rates. In particular, as the spray-solution flow rate increases through a pressure-swirl nozzle, additional energy is needed to achieve the same droplet (and particle) size. The additional energy is provided by applying increased atomization pressure.

During scale-up, the droplet drying rate is affected by two main factors. First, droplet density in the spray plume increases due to the higher spray-solution flow rate. Second, larger dryers are typically run in "recycle mode," where a fraction of the spray solvent is condensed from the drying gas before being recycled through the dryer during subsequent passes. Because not all the solvent is removed in the condenser, the solvent level in the drying gas can be higher for large-scale processes than for small-scale processes, which are typically run using a single pass of the drying gas.

To match the particle characteristics of SDDs made at large scale with those made at smaller scale, it is helpful to develop correlations of droplet size to solution viscosity, nozzle characteristics, and atomization parameters (eg, pressure). To complement these measurements, thermodynamic mass-energy-balance relationships can be modeled that take into account relative and absolute gas and spray-solution flow rates, process input and output temperatures, and relative residual-solvent vapor levels to define the thermodynamic operating space. This space can be assessed within the context of process equipment and plant constraints (eg, inlet drying-gas temperature).

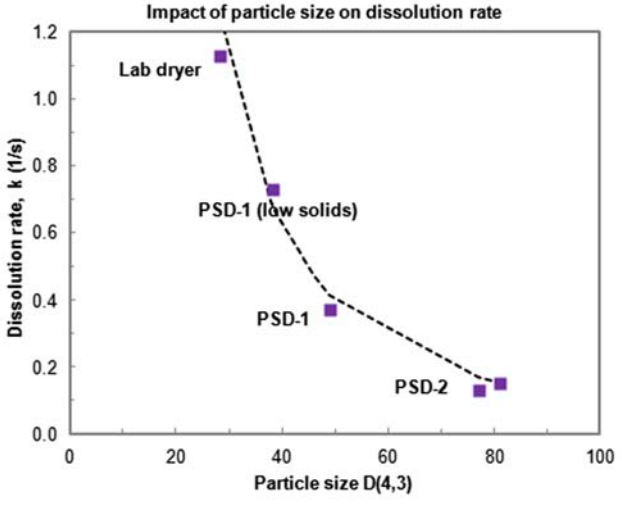


FIGURE 30.12 Effect of spray-drying scale and solids content on SDD particles, showing dissolution rate as a function of particle size.

30.5 SDD CHARACTERIZATION

A number of tools and techniques are available for characterizing SDD attributes including those shown in Table 30.4. These include tests for assessing physical characteristics, physical and chemical stability, powder properties, and in vitro dissolution performance.

30.5.1 Physical characteristics

Table 30.4 lists common analytical tools and techniques typically used to assess key physical properties of the SDD. These techniques range from particle properties (eg, size, morphology, density, wall thickness, hygroscopicity), the physical state of matter and stability (eg, dispersion homogeneity, crystallinity, propensity to crystallize, T_g, T_c).

30.5.2 Speciation testing

When dispersed in an aqueous environment, SDDs generally provide higher dissolved-drug concentrations than crystalline forms of the API because of the higher free energy of the amorphous form relative to the crystalline form. In addition, SDDs can produce drug–polymer colloids that can influence drug absorption. The mechanism and rate of SDD disintegration/dissolution and the relative abundance of various drug-containing species depends on (1) the drug properties; (2) the absolute and relative concentrations of the drug, polymers, and any other excipients in the SDD; and (3) the properties of the dissolution medium (eg, pH, the presence and abundance of bile-salt micelles).

Although the mechanism of drug release from SDDs can be complex, there are two mechanistic limits of SDD dissolution: (1) the rapid separation of drug- and polymer-rich phases, and (2) surface erosion, as shown in Fig. 30.13. In the former case, ingress of water induces a phase separation in which the SDD disintegrates into two or more phases, one of which tends to be drug–polymer colloids that are typically rich in drug relative to the initial SDD composition. These resulting drug-rich polymer colloids typically

TABLE 30.4 Analytical Techniques to Assess Performance of SDDs

SDD quality attributes	Typical analytical tools	Comments
Physical state and stability (presence of crystalline API, phase separation)	Dissolution tests	
	Powder X-Ray (PXRD)	Quantifying crystal formation and polymorphism. Sensitive to >5 wt% crystal formation can go <1 wt% crystal formation with high resolution PXRD
	Scanning Electron Microscope (SEM)	Evaluating particle morphology changes such as fusing and crystallization. Qualitative observation of crystals, but very sensitive
	Polarized Light Microscopy (PLM)	Presence of crystalline API. Qualitative observation of crystals, but very sensitive
	Modulated Differential Scanning Calorimetry (mDSC)	Evaluate SDD homogeneity, measure T_g, T_c, and T_m. Most sensitive for phase separation
	Isothermal Calorimetry—Thermal Activity Monitor (TAM)	Measure phase separation and crystallization kinetics. Very sensitive, but dependent on sample size
	Fourier Transform Infrared Spectroscopy (FT-IR)	Quantifying crystal formation and polymorphism
	Raman spectroscopy	Quantifying crystal formation and polymorphism
	Solid state NMR (ssNMR)	Quantifying crystal formation and polymorphism
Potency	High Performance Liquid Chromotograph (HPLC)	Method dependent
	In-line spectroscopy tools	
Chemical stability (impurities and degradation products)	HPLC	Impurity and method dependent. Additional analytical techniques (eg, NMR, LC-MS), may be used for identification
Dissolution	Evaluate ability of SDD stability samples to maintain supersaturation and sustainment	Estimated to be reproducible to 10%
Water content	Karl Fischer	50–100 ppm for coulometric KF
Residual solvent	Gas chromatography mass spectrometry (GC-MS)	Solvent and method dependent
	Quantitative Nuclear Magnetic Resonance (qNMR)	
Particle size	Laser light scattering particle size analysis	Optical or laser diffraction methods
	G3-Morphologi	
Surface area	Gas adsorption analysis	
Bulk density and tap density	Scott's volumeter	
True density	Helium pycnometer	
Morphology	G3 Morphologi	Morphology, particle structure, wall thickness
	SEM	
Hygroscopicity	Moisture-sorption	

provide rapid dissolution of drug. In the latter case, the polymer and drug on the SDD surface both dissolve, releasing more or less free drug and free polymer into the medium. The two mechanistic models suggested here are limits, and any particular SDD may dissolve by some combination of "erosion" and "disintegration."

Drug species typically provided by SDDs upon disintegration/dissolution include freely dissolved or solvated drug, drug in bile-salt micelles, and drug–polymer colloids. Significant numbers of drug–polymer colloids are typically formed only when using amphiphilic polymers, such as HPMCAS. These

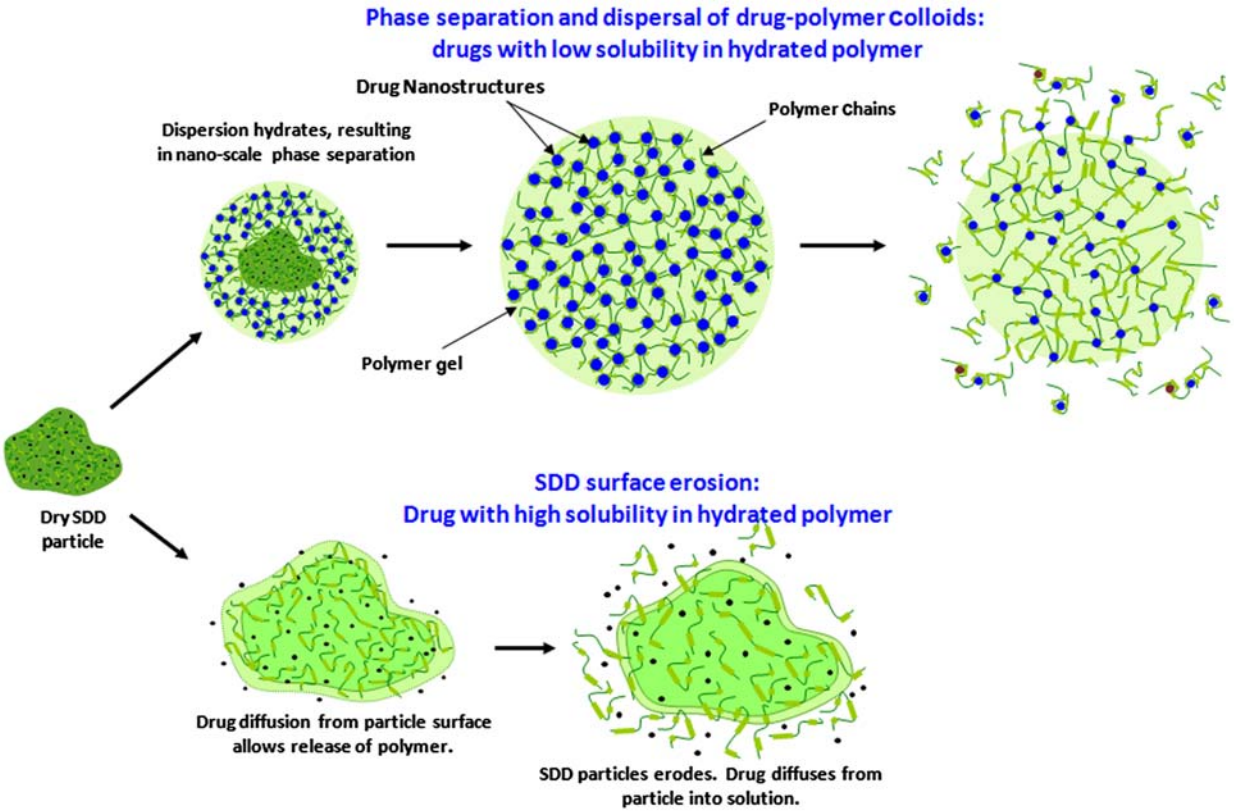


FIGURE 30.13 Two limits Of SDD dissolution mechanism: phase separation and dispersal of drug-rich polymer colloids (top) and surface erosion (bottom). *Source: Modified from Vodak DT, Morgen M. Design and Development of HPMCAS-Based Spray-Dried Dispersions. In: Shah N, Sandhu H, Choi DS, Chokshi H, Malick W, editors. Amorphous Solid Dispersions. New York: Springer; 2014. p. 303–22.*

colloids are high-energy, high-surface-area species that can enhance absorption by rapidly resupplying freely dissolved drug and drug in bile-salt micelles in the small intestine.[32] In addition, depending on their size, these drug–polymer colloids may enhance absorption in situations where absorption is limited by transport of drug through the unstirred mucus boundary layer of the epithelium by serving as drug carriers that transit this boundary layer. The SDD drug speciation and absorption mechanism is shown schematically in Fig. 30.14. Table 30.5 shows the types of tests that are performed to determine the drug species present.

30.5.3 In vitro dissolution testing

The general strategy behind solubilization technologies is to create a supersaturated solution and maintain the supersaturation long enough for drug absorption to take place in the intestine. Therefore, the main performance criteria for an SDD in in vitro testing are often rate, extent, and sustainment of dissolution in biorelevant aqueous media, such as simulated gastric fluid (SGF) or simulated intestinal fluid (SIF). Fig. 30.15 shows a representative dissolution-performance test plot showing the enhancement in drug concentration enabled by SDDs.[91] A number of in vitro dissolution tests have been used to assess performance, some of which are described next.

The solubilization performance of SDDs can be assessed in vitro using a wide variety of dissolution testing apparatuses and procedures. Dissolution testing often consists of, or is coupled with, additional analytical methods that lend insight into the SDD performance. Such methods are often aimed at determining the species of drug present, and ideally providing time-concentration profiles of the drug-containing species shown in Eq. (30.8). Where $[D_{total}]$ is total drug concentration, $[D_{free}]$ is free drug, $[D_{micelles}]$ is drug in bile-salt micelles, and $[DPN]$ is drug in drug/polymer nanoparticles.

$$[D_{total}] = [D_{free}] + [D_{micelles}] + [D_{PN}] \qquad (30.8)$$

An example of a performance-indicating in vitro test has been used by the authors is a small-scale dissolution test based on centrifugation to separate drug species. In this so-called "microcentrifuge test," SDD powder is added to a biorelevant medium in a

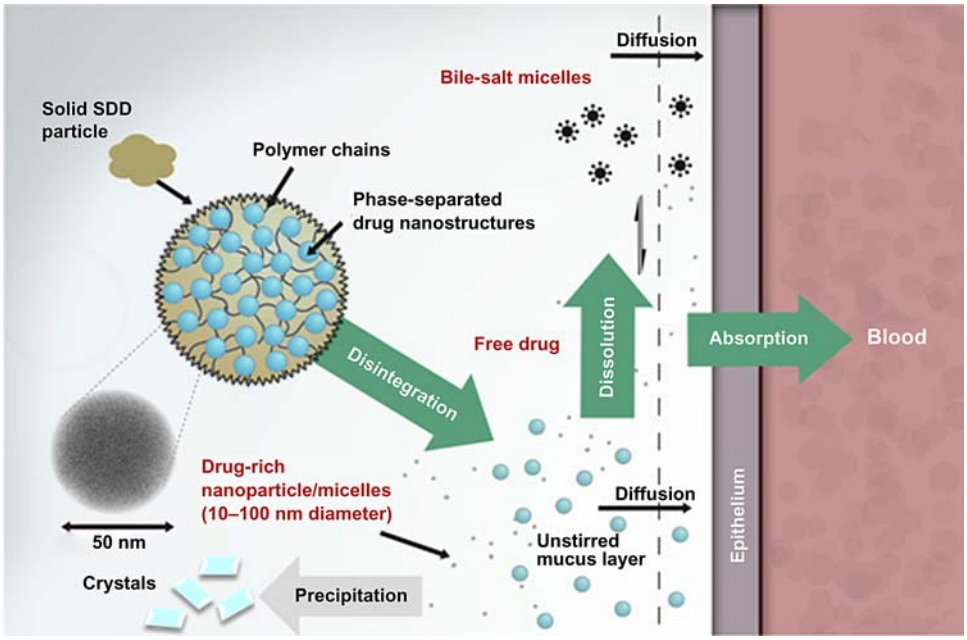


FIGURE 30.14 Schematic of SDD bioavailability enhancement, showing typical drug-containing species and absorption mechanisms.

TABLE 30.5 Drug Species During SDD Dissolution and Tools to Characterize Them

Drug species	Importance	In vitro determination
Undissolved drug or precipitate	• Source for high-energy drug • Evidence for conversion to low-energy form	• Microcentrifuge dissolution test • Collection and analysis of solids from solution (eg, potency, powder X-ray diffraction (PXRD))
Drug–polymer assemblies (ie, colloidal drug species or nanoaggregates)	• Relatively mobile source of high-energy drug	• Microcentrifuge dissolution test • Ultracentrifugation • Dynamic light-scattering (DLS) analysis
Drug in bile-salt micelles (or other micelles sourced from formulation)	• Highly mobile (mass transfer across unstirred boundary) • Critical in bioavailability of lipophilic compounds	• Partition coefficient determination • Microcentrifuge dissolution test • Membrane permeability test • Equilibrium solubility measurements • NMR and/or ultracentrifugation
Freely dissolved drug	• Highly mobile (mass transfer across unstirred boundary) • Driver for permeability • Driver for bile-salt micelle partitioning	• Membrane permeability test • Equilibrium solubility measurements • NMR and/or ultracentrifugation

microcentrifuge tube and vortexed to make a well-mixed suspension. The suspension is then periodically centrifuged, and the supernatant is sampled and analyzed for drug. The measured drug concentration typically consists of all three species shown in Eq. (30.8), all of which can play a role in absorption.

Ultracentrifugation can often be used to separate the drug in colloids from dissolved drug and drug in micelles. This process tends to pellet such colloidal material, depending on the colloids size and density. To confirm pelleting of colloidal material, dynamic light-scattering (DLS) can be used to detect any remaining colloids in the supernatant after micro- or ultracentrifugation.

NMR spectroscopy is another complementary tool that can be used to measure the concentration of dissolved drug and drug in micelles. Typically, drug in colloids is not visible by NMR due to the relatively slow tumbling of colloids in solution relative to the NMR frequency.

Any precipitate that forms or remains during such dissolution testing can be collected and analyzed by a variety of methods. Polarized light microscopy (PLM) can often be instructive in determining whether the

drug has crystallized during the test. Powder X-ray diffraction (PXRD) can be used as a more definitive technique to determine whether crystalline drug is present in the precipitate and to verify the crystalline form. To avoid crystallization of dissolved drug during drying of PXRD samples, wet precipitates can be collected and quickly frozen and lyophilized.

A second type of dissolution performance test that can be done at the small scale is a membrane-permeation test in which the SDD is placed in a biorelevant aqueous medium (eg, SIF) on one side of a semipermeable membrane and the concentration of drug in a permeant on the other side of the membrane is measured. The authors have used a polypropylene membrane treated to make the feed surface hydrophilic and used a low-volatility organic or an aqueous phase in which the drug is highly soluble as the sink or permeant.

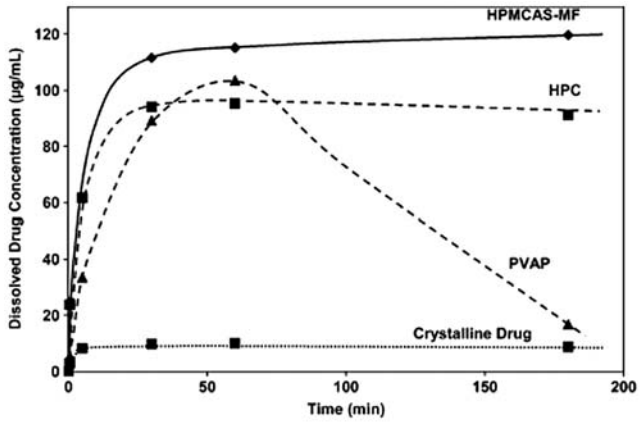


FIGURE 30.15 Dissolution performance of SDDs made with Compound 2 and various polymers, at 50% drug loading. Dissolution was carried out using the syringe dissolution test in MFDF at 37°C, with a 500 μg/mL total concentration (dissolved plus undissolved drug). *Source: From Curatolo W, Nightingale JA, Herbig SM. Utility of hydroxypropylmethylcellulose acetate succinate (HPMCAS) for initiation and maintenance of drug supersaturation in the GI milieu. Pharm Res 2009;26:1419–31. Published with permission of Springer.*

The dissolution test methods described previously are typically a sensitive measure of SDD performance and can be done at various stages of drug development for a variety of purposes. Early in formulation development, such testing can be done to screen various formulations to rank the performance of SDDs made with different polymers, drug loadings, or various other excipients, such as surfactants. At later stages of development, dissolution testing can be used to confirm the robustness of SDD manufacturing scale-up by comparing the performance of SDDs made at larger scales with those made early on at smaller scales. For example, the particle-size distribution can increase as the process is scaled from the smallest spray-dryers, producing larger SDD particles with reduced dissolution performance for some SDDs, as shown in Fig. 30.16.

Dissolution testing is often performed on an SDD after accelerated aging studies (eg, storage at high temperatures and humidities) to predict the long-term physical stability of the formulation. The dissolution performance is typically very sensitive to fusing of SDD particles, drug crystallization, or amorphous phase separation. Finally, the speciation data obtained from the comprehensive dissolution studies described previously can be used as inputs to mechanistic physiologically based pharmacokinetic models to predict oral absorption of various SDD formulations relative to each other and other formulation approaches.

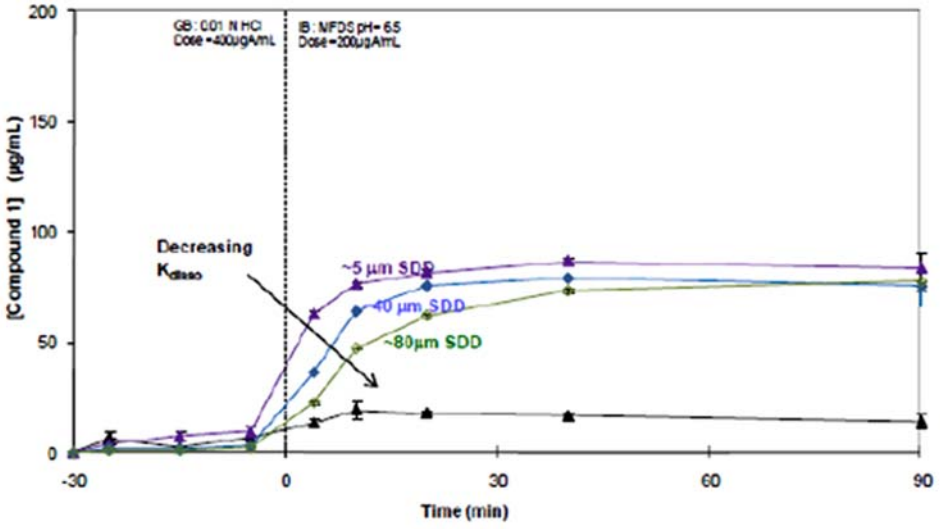


FIGURE 30.16 Dissolution test results showing the impact of larger particle sizes on dissolution performance in a SGF/SIF transfer test. Conditions: SGF = 0.01 N HCl at 400 μgA/mL, SIF = DEFINE at 200 μgA/mL.

TABLE 30.6 Purposes of Typical Dissolution Tests, Uses, and Outputs

Purpose	When	Scope	Output
Formulation screening	SDD feasibility	Dissolution in SGF-SIF and SIF, free drug and dynamic light scattering	Rank ordering of SDD performance
Process robustness	Scale-up or change in process	Differences in rate, C_{max}, or AUC with process/scale change	Performance risk with increasing scale, particle size, or process parameters
Physical stability	Advanced SDD screening or IND supporting stability	Differences in rate, C_{max}, or AUC relative to SDD physical stability during accelerated aging	Performance risk related to physical or chemical changes with long-term storage
Mechanistic absorption model	SDD feasibility	Dissolved, micellar, and colloidal drug levels SGF-SIF vs. SIF alone	Time-concentration profile of drug species that are inputs for oral absorption model—level of sophistication fit for purpose

SGF—Simulated Gastric Fluid.
SIF—Simulated Intestinal Fluid.
SGF-SIF—A two step transfer dissolution method.
C_{max}—Maximum concentration.
AUC—Area Under Curve.

Table 30.6 summarizes the various applications for dissolution testing and the key output from such tests.

30.5.4 Physical/chemical stability during storage as suspension or powder

The physical and chemical stability of the SDD during short- and long-term storage is important. As high-energy solid forms, SDDs can be prone to physical instability, including amorphous phase separation or drug crystallization. Likewise, due to the lack of crystalline order, the drug in an SDD can sometimes be more prone to chemical degradation than when in crystalline form. The increased susceptibility of amorphous drug forms to chemical degradation makes it important to assess the compatibility of the API with the SDD excipients, particularly the polymer. In addition, it may be necessary to use additives, such as antioxidants, to protect APIs that are susceptible to specific mechanisms of degradation.

The typical two modes of physical instability—amorphous phase separation and drug crystallization—are shown schematically in Fig. 30.17. Drug crystallization often follows after separation of amorphous domains has occurred because once amorphous domains of pure or enriched drug phases have formed, less drug mobility is needed for crystallization to occur.

Phase separation in SDDs is driven by molecular mobility of the polymer matrix and drug within the matrix. One indicator of the molecular mobility is the difference between T_s (Storage Temperature) and T_g. As the T_s approaches the T_g of the SDD, mobility increases, viscosity decreases, and the time to phase separation, including crystallization decreases.

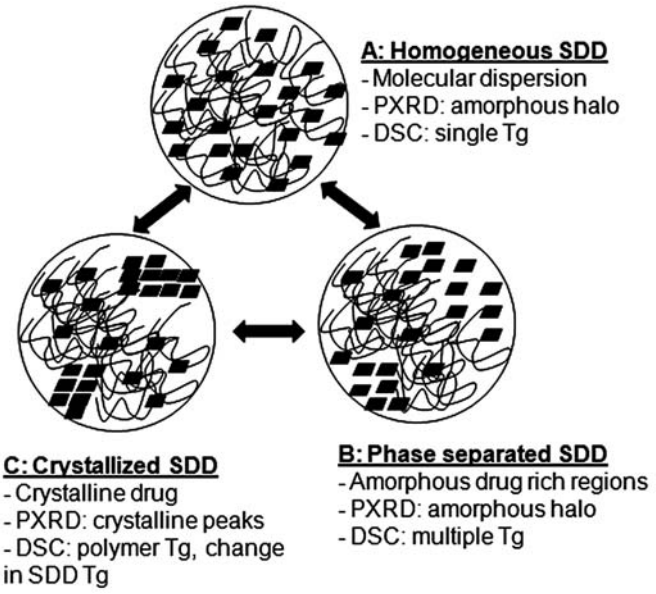


FIGURE 30.17 Possible structures of SDDs where curvy lines and parallelogram symbol represent polymer and drug, respectively. A is homogeneous molecular dispersion, B is phase-separated SDD containing amorphous-rich drug region, and C is crystallized SDD containing crystalline drug. *Source: Adapted from*[14] *Huang Y, Dai WG. Fundamental aspects of solid dispersion technology for poorly soluble drugs. Acta Pharm Sin B 2014; 4:18–25.*

Since water can plasticize the SDD lowering the T_g, it is useful to understand SDD physical stability as a function of RH.

Early in a SDD-development program physical stability predictions of the SDD are often based on T_g versus RH curves (Fig. 30.18). It is assumed that if T_g values are significantly higher than the anticipated

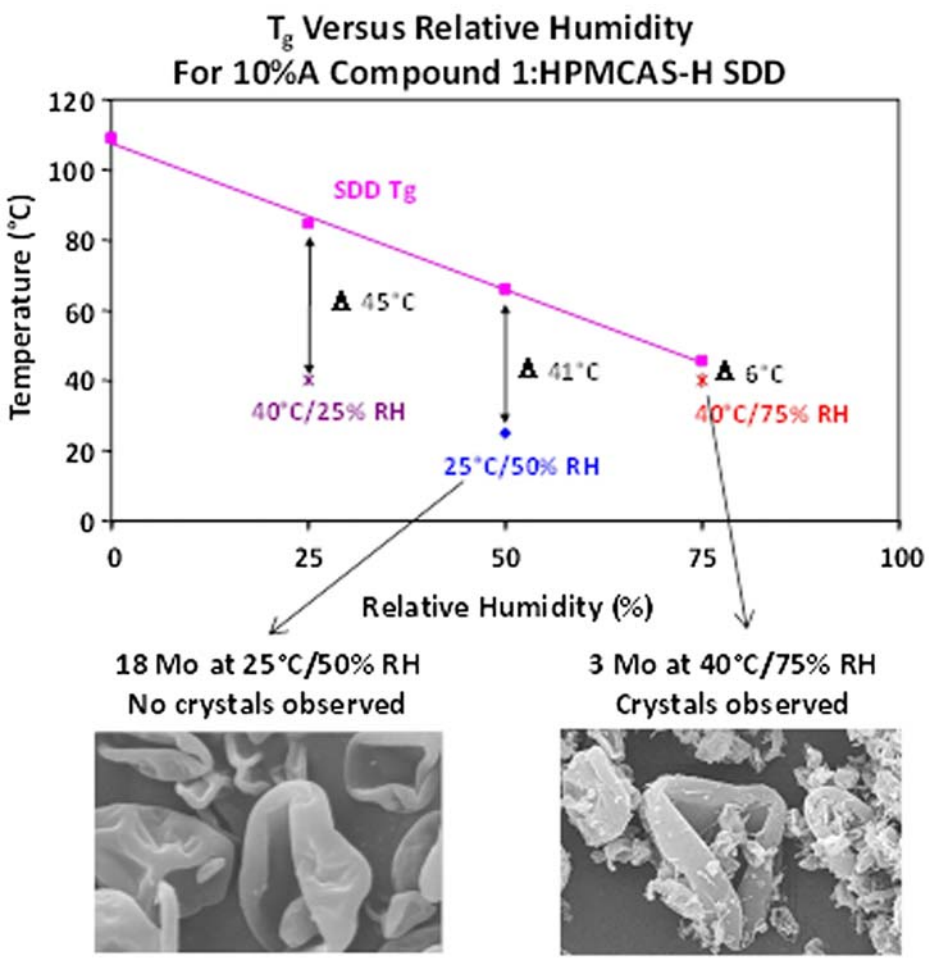


FIGURE 30.18 T_g versus RH for an example 10%A HPMCAS SDD.

T_s, physical stability will be acceptable. For SDDs with a T_g close to or below the anticipated T_s, physical stability could become an issue. Historically, a good rule of thumb was that storage at 50°C below the T_g should result in good long-term physical stability.[38] This turned out to be a conservative metric with regards to ensuring a 2-year shelf life for SDD pharmaceutical products. In the example shown in Fig. 30.18, a 10%A HPMCAS SDD demonstrated good physical stability when stored for 18 months at 25°C/50% RH when the T_g was only 41°C above the T_s, whereas when stored at 40°C/75% RH (at a T_g 6°C above T_s), the SDD crystallized within 3 months.

Fig. 30.19 shows the fraction of SDDs that were found to be physically stable at 6- and 13-week time points as a function of $T_g - T_s$. The data, taken from a variety of development programs, suggest that when the T_g is well above T_s, the SDD has good stability. As the T_g drops relative to T_s, the risk of physical instability increases. In this instance, physical stability was defined as no significant change in thermal characteristics (by differential scanning calorimetry (DSC)), morphology (by SEM), or in vitro dissolution performance.

30.6 DOSAGE FORM CONSIDERATIONS

SDDs are amenable to administration in a wide range of dosage forms, including as a suspension or in a solid dosage form, such as tablets and capsules.

30.6.1 Aqueous SDD suspension formulations

It is becoming increasingly common to use aqueous SDD suspensions to support (1) animal safety and efficacy studies, where dose multiples are significantly higher than clinical doses; and (2) exploratory clinical studies as powder-in-bottle formulations. Typically, a suspending agent such as methyl cellulose is used to increase the viscosity of the SDD suspension to improve homogeneity and stability. Buffers, additional polymers, surfactants, and other excipients may be used to improve physical stability of the aqueous suspension.[92] Suspension formulations are generally prepared and administered within a short period of time, depending on physical and chemical stability. In many cases, it is possible to achieve suspension stability of

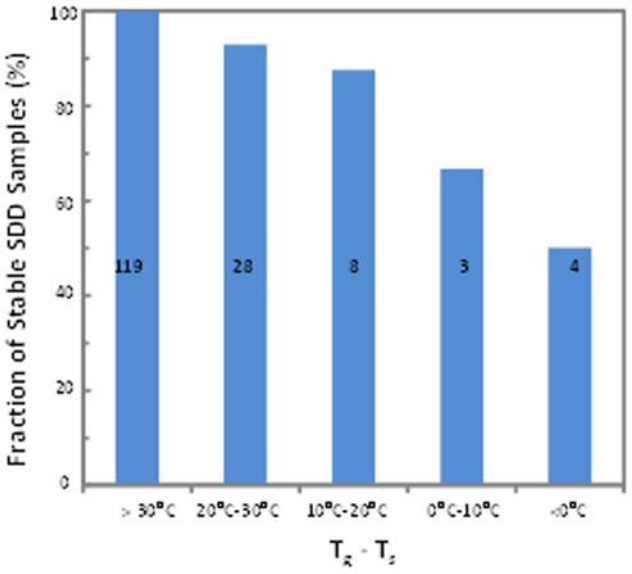


FIGURE 30.19 Fraction of SDDs that are Stable after storage for up to 13 weeks as a function of $T_g - T_s$. Key: Numbers in columns indicate the number of stability samples tested at 6 or 13 weeks for compounds transferred into good manufacturing practice (GMP) production. Stability is defined as no change in dissolution performance, thermal characteristics (by DSC) or morphology (based on SEM analysis) from that of initial samples.

TABLE 30.7 Typical Syringable Volume and Dose for SDD Formulations Prepared Using HPMCAS-M and 0.5% W/W Methocel E4M Premium as a Suspending Agent

SDD drug loading (%Active)	Maximum syringable concentration (mgA/mL)	Deliverable dose (mg/kg) at 10 mL/kg
10	17	167
25	50	500
33	75	749
40	100	1000
50	150	1500
75	450	4500

days to a week. Typical doses achievable in safety studies are shown as a function of API loading in the SDD provided in Table 30.7.

30.6.2 Solid dosage forms

Traditionally, immediate-release tablets are the most common solid dosage forms for SDDs, although capsule formulations have also been used. The dosage form should allow fast disintegration and release of native SDD particles in the stomach or small intestine. Designing the dosage form to quickly and uniformly disperse the SDD avoids the gelling and subsequent slow disintegration and release that is typically undesirable for bioavailability-enhancement applications.

For tablet formulations, the SDD is typically dry granulated (using roller compaction) or directly compressed. Wet granulation is usually avoided due to the risk of process-induced physical instability.

Some of the key considerations during tablet formulation development include: rapid disintegration, drug-excipient compatibility, physical and chemical stability during processing and storage, drug loading and maximum achievable dose, and choice of granulation/compression process.

One of the important considerations during solid-dosage-form development is the maximum unit strength, which is governed by the SDD loading in the solid dosage form and the API loading in the SDD (Fig. 30.20). For low-dose compounds, the strategy is generally to have a lower API loading in the SDD and a higher SDD loading in the tablet. For high-dose compounds, API loading in the SDD and the SDD loading in the tablet both must be maximized, which can result in performance and manufacturability risk for the SDD and tablets. Some of the constraints to consider are the maximum tablet size allowed, the effect of SDD loading in tablet on dissolution, and the effect of SDD material properties on tableting.

30.6.3 Effect of formulation and process on performance (physical, chemical, dissolution)

Although the impact of drug-product processing on crystalline API phase transformation is well-studied, there is limited literature describing the effects of solid-dosage-form processing on SDD stability. It is reasonable to expect that the physical stability of the SDD may be affected by the formulation composition and manufacturing process. Leane et al. recently demonstrated a marked difference in physical stability of ibipinabant SDD tablets (20/75/5 ibipinabant/PVP K30/SLS SDD, $T_g = 130$ºC, 5% w/w SDD in tablet formulation) in the presence of different fillers. Microcrystalline cellulose (MCC) provided the best stability profile and mannitol provided the worst, as shown in Fig. 30.21.[62] As expected, the physical instability (ie, crystallization of API in SDD) resulted in a corresponding decrease in dissolution rate. The authors attributed the higher stability achieved with MCC to the cushioning it provided during compaction, preventing fracturing of the SDD particles. Dhumal et al. had previously shown that the tablets manufactured with amorphous coprecipitate of celecoxib with PVP and carrageenan had better stability than tablets manufactured with amorphous celecoxib and an SDD

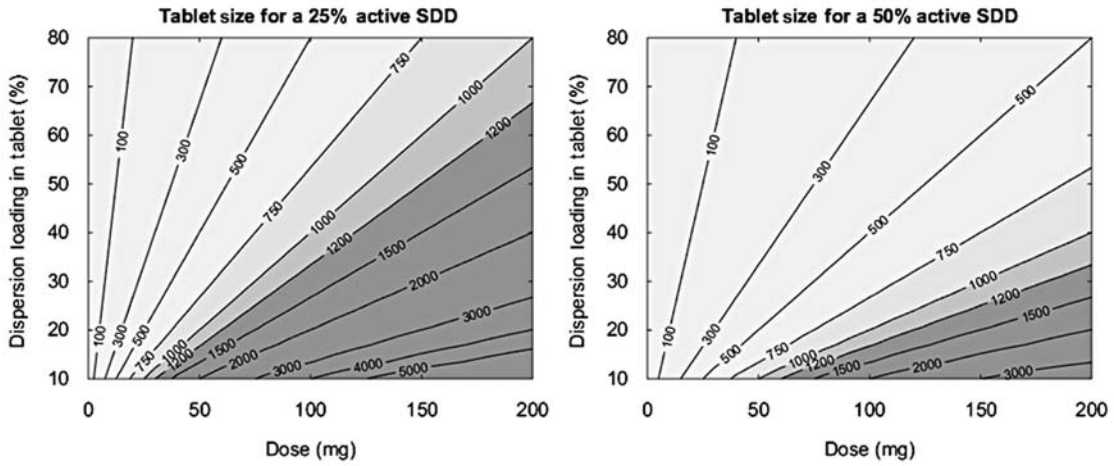


FIGURE 30.20 Effect of SDD loading in the tablet and dose on tablet size for 25%A and 50%A SDDs.

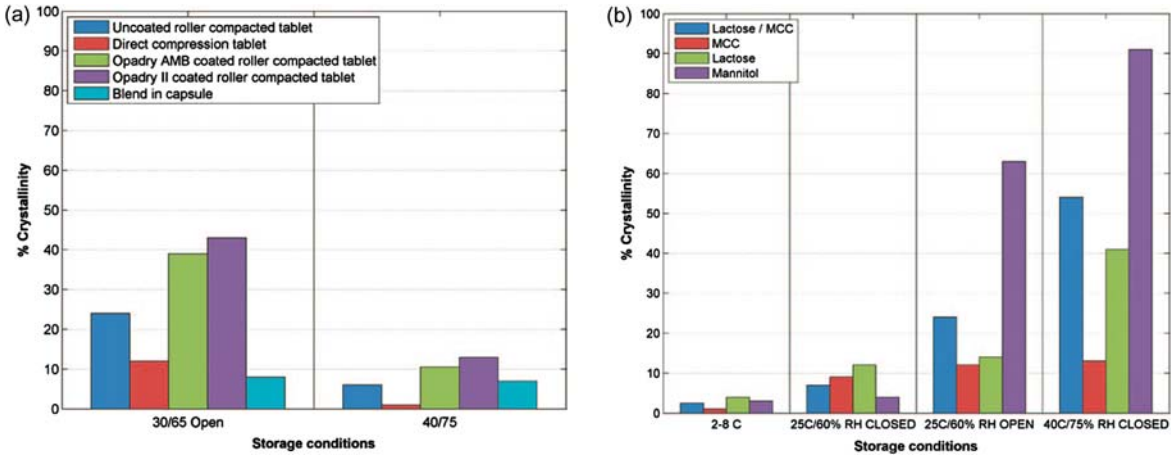


FIGURE 30.21 Raman results for formulations containing different excipients (a) and different processing conditions (b) following storage under controlled conditions for 3 months. *Source: From Leane MM, Sinclair W, Qian F, Haddadin R, Brown A, Tobyn M, et al. Formulation and process design for a solid dosage form containing a spray-dried amorphous dispersion of ibipinabant. Pharm Dev Technol 2013;18:359–66. Published with permission from Taylor & Francis Ltd., http://www.informaworld.com.*

of celecoxib with PVP.[93] The authors attributed the improved physical stability of the coprecipitate to the cushioning provided by the viscoelastic polymer carageenan and hydrogen-bonding interaction between celecoxib and PVP.

Crystallization of amorphous solids can involve bulk and surface crystallization. Surface crystal growth has been observed to be orders of magnitude faster than bulk crystal growth due to the enhanced surface mobility and energy of the drug molecules. An understanding of surface–surface interactions and the impact of excipients on surface crystallization is helpful in designing the SDD formulations because SDDs have an intimate contact with other excipients within a solid dosage form. Zhang et al. recently reported a simple microscopic method for studying the effect of common tablet excipients on the surface crystallization of amorphous drugs.[94] In this method, amorphous samples of the drugs were covered by excipients and stored in controlled environments and rates of surface crystal nucleation and growth were measured in real-time.

When considering the choice of excipients, excess manipulation (eg, compression, milling) during manufacture of the solid dosage form and exposure of the solid dosage form to moisture and heat during the coating process can lead to physical instability of SDD formulations.

30.7 CONCLUDING REMARKS

Amorphous SDDs are ideally molecularly dispersed mixtures of drug, polymers, and optionally, other

excipients (eg, surfactants, stabilizers). SDDs stabilize amorphous drug, resulting in reduced crystallization by inhibiting molecular mobility and reducing nucleation. SDDs typically provide higher apparent solubility and faster dissolution than bulk drug crystals, affording superior bioperformance for poorly soluble drugs. Spray-drying is a robust and scalable technology that has been successfully applied to a number of commercial products including Kalydeco (Ivacaftor), Incivek (Telaprevir), Intelence (Etravirine), and Zortress (Everolimus). On the basis of a literature survey, Newman et al. estimated that 82% of the amorphous dispersions showed improved bioavailability over a reference formulation.[13] SDDs of more than 500 different drugs showed 1.5- to 100-fold higher absorption than crystalline drug in various animal models. In humans, SDDs of 65 different drugs showed at least twofold higher relative bioavailability than the poorly absorbed control formulation. A high success rate in enhancing oral absorption of poorly absorbed compounds, versatility, good fundamental understanding, and availability of spray dryers that can make SDDs from milligram to metric ton quantities has made SDDs a common approach for developing of compounds in preclinical and clinical development.

Successful development of SDDs requires careful consideration of CQAs (eg, physical and chemical stability, in vitro and in vivo performance, residual solvents, and material properties for downstream processing) that can impact the TPP. It also requires a good understanding of the formulation (eg, API and polymer chemistry, API:polymer ratio, excipient, and solvent) and process parameters (eg, droplet size, drying rate, and secondary drying) that define the SDD attributes.

References

1. Kawabata Y, Wada K, Nakatani M, Yamada S, Onoue S. Formulation design for poorly water-soluble drugs based on biopharmaceutics classification system: basic approaches and practical applications. *Int J Pharm* 2011;**420**:1–10.
2. Buckley ST, Frank KJ, Fricker G, Brandl M. Biopharmaceutical classification of poorly soluble drugs with respect to "enabling formulations". *Eur J Pharm Sci* 2013;**50**:8–16.
3. Singh A, Worku ZA, Van den Mooter G. Oral formulation strategies to improve solubility of poorly water-soluble drugs. *Expert Opin Drug Deliv* 2011;**8**:1361–78.
4. Williams HD, Trevaskis NL, Charman SA, Shanker RM, Charman WN, Pouton CW, et al. Strategies to address low drug solubility in discovery and development. *Pharmacol Rev* 2013;**65**:315–499.
5. Rabinow BE. Nanosuspensions in drug delivery. *Nat Rev Drug Discov* 2004;**3**:785–96.
6. Frank KJ, Rosenblatt KM, Westedt U, Holig P, Rosenberg J, Magerlein M, et al. Amorphous solid dispersion enhances permeation of poorly soluble ABT-102: true supersaturation vs. apparent solubility enhancement. *Int J Pharm* 2012;**437**:288–93.
7. Vo CL, Park C, Lee BJ. Current trends and future perspectives of solid dispersions containing poorly water-soluble drugs. *Eur J Pharm Biopharm* 2013;**85**:799–813.
8. Brewster ME, Loftsson T. Cyclodextrins as pharmaceutical solubilizers. *Adv Drug Deliv Rev* 2007;**59**:645–66.
9. Chen XQ, Gudmundsson OS, Hageman MJ. Application of lipid-based formulations in drug discovery. *J Med Chem* 2012;**55**:7945–56.
10. Serajuddin AT. Salt formation to improve drug solubility. *Adv Drug Deliv Rev* 2007;**59**:603–16.
11. Vig BS, Huttunen KM, Laine K, Rautio J. Amino acids as promoieties in prodrug design and development. *Adv Drug Deliv Rev* 2013;**65**:1370–85.
12. Zhang GZ, Zhou D. Crystalline and amorphous solids. In: Qiu Y, Chen Y, Zhang GZ, Liu L, Porter W, editors. *Developing solid oral dosage forms: pharmaceutical theory & practice*. 1st ed. Amsterdam, London, Boston, MA: Academic Press; 2009. p. 25–60.
13. Newman A, Knipp G, Zografi G. Assessing the performance of amorphous solid dispersions. *J Pharm Sci* 2012;**101**:1355–77.
14. Huang Y, Dai WG. Fundamental aspects of solid dispersion technology for poorly soluble drugs. *Acta Pharm Sin B* 2014;**4**:18–25.
15. Stella VJ, Nti-Addae KW. Prodrug strategies to overcome poor water solubility. *Adv Drug Deliv Rev* 2007;**59**:677–94.
16. Khadka P, Ro J, Kim H, Kim I, Kim JT, Kim H, et al. Pharmaceutical particle technologies: an approach to improve drug solubility, dissolution and bioavailability. *Asian J Pharm Sci* 2014;**9**:304–16.
17. Oh DM, Curl RL, Amidon GL. Estimating the fraction dose absorbed from suspensions of poorly soluble compounds in humans: a mathematical model. *Pharm Res* 1993;**10**:264–70.
18. Chen XQ, Stefanski K, Shen H, Huang C, Caporuscio C, Yang W, et al. Oral delivery of highly lipophilic poorly water-soluble drugs: spray-dried dispersions to improve oral absorption and enable high-dose toxicology studies of a P2Y1 antagonist. *J Pharm Sci* 2014;**103**:3924–31.
19. Bhatnagar P, Dhote V, Mahajan SC, Mishra PK, Mishra DK. Solid dispersion in pharmaceutical drug development: from basics to clinical applications. *Curr Drug Deliv* 2014;**11**:155–71.
20. Leuner C, Dressman J. Improving drug solubility for oral delivery using solid dispersions. *Eur J Pharm Biopharm* 2000;**50**:47–60.
21. Dobry DE, Settell DM, Baumann JM. *Spray drying and scale-up. Pharmaceutical sciences encyclopedia*, vol. 9. New York, NY: John Wiley & Sons, Inc; 2010. p. 1–26.
22. Verreck G, Six K, Van den Mooter G, Baert L, Peeters J, Brewster ME. Characterization of solid dispersions of itraconazole and hydroxypropylmethylcellulose prepared by melt extrusion--Part I. *Int J Pharm* 2003;**251**:165–74.
23. Shah N, Sandhu H, Phuapradit W, Pinal R, Iyer R, Albano A, et al. Development of novel microprecipitated bulk powder (MBP) technology for manufacturing stable amorphous formulations of poorly soluble drugs. *Int J Pharm* 2012;**438**:53–60.
24. Betageri GV, Makarla KR. Enhancement of dissolution of glyburide by solid dispersion and lyophilization techniques. *Int J Pharm* 1995;**126**:155–60.
25. Dinunzio JC, Brough C, Hughey JR, Miller DA, Williams RO, McGinity JW. Fusion production of solid dispersions containing a heat-sensitive active ingredient by hot melt extrusion and Kinetisol dispersing. *Eur J Pharm Biopharm* 2010;**74**:340–51.
26. Douroumis D. *Hot melt extrusion: pharmaceutical applications*. Chichester: Wiley; 2012.

27. Lambert WJ. Considerations in developing a target product profile for parenteral pharmaceutical products. *AAPS PharmSciTech* 2010;**11**:1476–81.
28. Guidance for Industry. Q8(R2) Pharmaceutical development. In: US Food and Drug Administration editor; 2009.
29. Paudel A, Worku ZA, Meeus J, Guns S, Van den Mooter G. Manufacturing of solid dispersions of poorly water soluble drugs by spray drying: formulation and process considerations. *Int J Pharm* 2013;**453**:253–84.
30. Sonje VM, Kumar L, Puri V, Kohli G, Kaushal AM, Bansal AK. Effect of counterions on the properties of amorphous atorvastatin salts. *Eur J Pharm Sci* 2011;**44**:462–70.
31. Tong P, Taylor LS, Zografi G. Influence of alkali metal counterions on the glass transition temperature of amorphous indomethacin salts. *Pharm Res* 2002;**19**:649–54.
32. Friesen DT, Shanker R, Crew M, Smithey DT, Curatolo WJ, Nightingale JA. Hydroxypropyl methylcellulose acetate succinate-based spray-dried dispersions: an overview. *Mol Pharm* 2008;**5**:1003–19.
33. Trasi NS, Baird JA, Kestur US, Taylor LS. Factors influencing crystal growth rates from undercooled liquids of pharmaceutical compounds. *J Phys Chem B* 2014;**118**:9974–82.
34. Tang XC, Pikal MJ, Taylor LS. The effect of temperature on hydrogen bonding in crystalline and amorphous phases in dihydropyrine calcium channel blockers. *Pharm Res* 2002;**19**:484–90.
35. Baird JA, Van Eerdenbrugh B, Taylor LS. A classification system to assess the crystallization tendency of organic molecules from undercooled melts. *J Pharm Sci* 2010;**99**:3787–806.
36. Van Eerdenbrugh B, Baird JA, Taylor LS. Crystallization tendency of active pharmaceutical ingredients following rapid solvent evaporation--classification and comparison with crystallization tendency from undercooled melts. *J Pharm Sci* 2010;**99**:3826–38.
37. Masuda K, Tabata S, Sakata Y, Hayase T, Yonemochi E, Terada K. Comparison of molecular mobility in the glassy state between amorphous indomethacin and salicin based on spin-lattice relaxation times. *Pharm Res* 2005;**22**:797–805.
38. Hancock BC, Shamblin SL, Zografi G. Molecular mobility of amorphous pharmaceutical solids below their glass transition temperatures. *Pharm Res* 1995;**12**:799–806.
39. Ghebremeskel AN, Vemavarapu C, Lodaya M. Use of surfactants as plasticizers in preparing solid dispersions of poorly soluble API: selection of polymer–surfactant combinations using solubility parameters and testing the processability. *Int J Pharm* 2007;**328**:119–29.
40. Towler CS, Li T, Wikstrom H, Remick DM, Sanchez-Felix MV, Taylor LS. An investigation into the influence of counterion on the properties of some amorphous organic salts. *Mol Pharm* 2008;**5**:946–55.
41. Li J, Zhao J, Tao L, Wang J, Waknis V, Pan D, et al. The effect of polymeric excipients on the physical properties and performance of amorphous dispersions: part I, free volume and glass transition. *Pharm Res* 2015;**32**:500–15.
42. Qian F, Wang J, Hartley R, Tao J, Haddadin R, Mathias N, et al. Solution behavior of PVP-VA and HPMC-AS-based amorphous solid dispersions and their bioavailability implications. *Pharm Res* 2012;**29**:2765–76.
43. Teja SB, Patil SP, Shete G, Patel S, Bansal AK. Drug-excipient behavior in polymeric amorphous solid dispersions. *J Excipients and Food Chem.* 2013;**4**:70–94.
44. Qian F, Huang J, Hussain MA. Drug-polymer solubility and miscibility: stability consideration and practical challenges in amorphous solid dispersion development. *J Pharm Sci* 2010;**99**:2941–7.
45. Paudel A, Van Humbeeck J, Van den Mooter G. Theoretical and experimental investigation on the solid solubility and miscibility of naproxen in poly(vinylpyrrolidone). *Mol Pharm* 2010;**7**:1133–48.
46. Weuts I, Kempen D, Verreck G, Peeters J, Brewster M, Blaton N, et al. Salt formation in solid dispersions consisting of polyacrylic acid as a carrier and three basic model compounds resulting in very high glass transition temperatures and constant dissolution properties upon storage. *Eur J Pharm Sci* 2005;**25**:387–93.
47. Yoo SU, Krill SL, Wang Z, Telang C. Miscibility/stability considerations in binary solid dispersion systems composed of functional excipients towards the design of multi-component amorphous systems. *J Pharm Sci* 2009;**98**:4711–23.
48. Taylor LS, Zografi G. Spectroscopic characterization of interactions between PVP and indomethacin in amorphous molecular dispersions. *Pharm Res* 1997;**14**:1691–8.
49. Fox TG. Influence of diluent and of copolymer composition on the glass temperature of a polymer system. *Bull Am Phys Soc* 1956;**1**:123.
50. Gordon M, Taylor LS. Ideal copolymers and the second-order transitions of synthetic rubbers. I. Non-crystalline copolymers. *J Appl Chem* 1952;**2**:493–500.
51. Konno H, Handa T, Alonzo DE, Taylor LS. Effect of polymer type on the dissolution profile of amorphous solid dispersions containing felodipine. *Eur J Pharm Biopharm* 2008;**70**:493–9.
52. Alonzo D, Zhang GZ, Zhou D, Gao Y, Taylor L. Understanding the behavior of amorphous pharmaceutical systems during dissolution. *Pharm Res* 2010;**27**:608–18.
53. Guzman HR, Tawa M, Zhang Z, Ratanabanangkoon P, Shaw P, Gardner CR, et al. Combined use of crystalline salt forms and precipitation inhibitors to improve oral absorption of celecoxib from solid oral formulations. *J Pharm Sci* 2007;**96**:2686–702.
54. Zecevic DE, Meier R, Daniels R, Wagner KG. Site specific solubility improvement using solid dispersions of HPMC-AS/HPC SSL--mixtures. *Eur J Pharm Biopharm* 2014;**87**:264–70.
55. Miller DA, DiNunzio JC, Yang W, McGinity JW, Williams III RO. Enhanced in vivo absorption of itraconazole via stabilization of supersaturation following acidic-to-neutral pH transition. *Drug Dev Ind Pharm* 2008;**34**:890–902.
56. Patterson A, Ferreira AP, Banks E, Skeene K, Clarke G, Nicholson S, et al. Modelling drug degradation in a spray dried polymer dispersion using a modified Arrhenius equation. *Int J Pharm* 2014;**478**:348–60.
57. Gupta P, Bansal AK. Ternary amorphous composites of celecoxib, poly(vinyl pyrrolidone) and meglumine with enhanced solubility. *Pharmazie* 2005;**60**:830–6.
58. Gupta P, Bansal AK. Spray drying for generation of a ternary amorphous system of celecoxib, PVP, and meglumine. *Pharm Dev Technol* 2005;**10**:273–81.
59. Pokharkar VB, Mandpe LP, Padamwar MN, Ambike AA, Mahadik KR, Paradkar A. Development, characterization and stabilization of amorphous form of a low T_g drug. *Powder Technol* 2006;**167**:20–5.
60. Takeuchi H, Nagira S, Yamamoto H, Kawashima Y. Solid dispersion particles of amorphous indomethacin with fine porous silica particles by using spray-drying method. *Int J Pharm* 2005;**293**:155–64.
61. Koester LS, Mayorga P, Bassani VL. Carbamazepine/betaCD/HPMC solid dispersions. I. Influence of the spray-drying process and betaCD/HPMC on the drug dissolution profile. *Drug Dev Ind Pharm* 2003;**29**:139–44.
62. Leane MM, Sinclair W, Qian F, Haddadin R, Brown A, Tobyn M, et al. Formulation and process design for a solid dosage form containing a spray-dried amorphous dispersion of ibipinabant. *Pharm Dev Technol* 2013;**18**:359–66.

63. Yan Y-D, Sung JH, Kim KK, Kim DW, Kim JO, Lee B-J, et al. Novel valsartan-loaded solid dispersion with enhanced bioavailability and no crystalline changes. *Int J Pharm* 2012;**422**:202–10.
64. Janssens S, Nagels S, Armas HNd, D'Autry W, Van Schepdael A, Van den Mooter G. Formulation and characterization of ternary solid dispersions made up of Itraconazole and two excipients, TPGS 1000 and PVPVA 64, that were selected based on a supersaturation screening study. *Eur J Pharm Biopharm* 2008;**69**:158–66.
65. Cho Y, Ha ES, Baek IH, Kim MS, Cho CW, Hwang SJ. Enhanced supersaturation and oral absorption of sirolimus using an amorphous solid dispersion based on Eudragit® e. *Molecules* 2015;**20**:9496–509.
66. Goddeeris C, Willems T, Van den Mooter G. Formulation of fast disintegrating tablets of ternary solid dispersions consisting of TPGS 1000 and HPMC 2910 or PVPVA 64 to improve the dissolution of the anti-HIV drug UC 781. *Eur J Pharm Sci* 2008;**34**:293–302.
67. Guo Y, Luo J, Tan S, Otieno BO, Zhang Z. The applications of Vitamin E TPGS in drug delivery. *Eur J Pharm Sci* 2013;**49**:175–86.
68. Gamble JF, Ferreira AP, Tobyn M, DiMemmo L, Martin K, Mathias N, et al. Application of imaging based tools for the characterisation of hollow spray dried amorphous dispersion particles. *Int J Pharm* 2014;**465**:210–17.
69. ICHQ3(R5). Impurities guideline for residual solvents, In: International Conference on Harmonisation, Geneva; 2011.
70. Paudel A, Van den Mooter G. Influence of solvent composition on the miscibility and physical stability of naproxen/PVP K 25 solid dispersions prepared by cosolvent spray-drying. *Pharm Res* 2012;**29**:251–70.
71. Kamlet MJ, Abboud JLM, Abraham MH, Taft RW. Linear solvation energy relationships. 23. A comprehensive collection of the solvatochromic parameters, .pi.*, .alpha., and .beta., and some methods for simplifying the generalized solvatochromic equation. *J Org Chem* 1983;**48**:2877–87.
72. Tung HH, Tabora J, Variankaval N, Bakken D, Chen CC. Prediction of pharmaceutical solubility Via NRTL-SAC and COSMO-SAC. *J Pharm Sci* 2008;**97**:1813–20.
73. Al-Obaidi H, Brocchini S, Buckton G. Anomalous properties of spray dried solid dispersions. *J Pharm Sci* 2009;**98**:4724–37.
74. Chen R, Ilasi N, Sekulic SS. Absolute molecular weight determination of hypromellose acetate succinate by size exclusion chromatography: use of a multi angle laser light scattering detector and a mixed solvent. *J Pharm Biomed Anal* 2011;**56**:743–8.
75. Vehring R. Pharmaceutical particle engineering via spray drying. *Pharm Res* 2008;**25**:999–1022.
76. DuBose D, Settell D, Bennette N, Broadbent A. Spray-dried dispersions—developing process control strategies for the manufacture of spray-dried dispersions. *Drug Dev Deliv* 2015;**15**:75–81.
77. Quirk E, Doggett A, Bretnall A. Determination of residual acetone and acetone related impurities in drug product intermediates prepared as spray dried dispersions (SDD) using gas chromatography with headspace autosampling (GCHS). *J Pharm Biomed Anal* 2014;**96**:37–44.
78. Yue H, Nicholson SJ, Young JD, Hsieh D, Ketner RJ, Hall RG, et al. Development of a control strategy for benzene impurity in HPMCAS-stabilized spray-dried dispersion drug products using a science-based and risk-based approach. *Pharm Res* 2015;**32**:2636–48.
79. Friesen DT, Newbold DD, Baumann JM, Dubose DB, Millard DL. Spray-drying process. Google Patents; 2012.
80. Vu H, Aguliar G. High-speed internal nozzle flow visualization of flashing jets. In: *11th triennial international annual conference on liquid atomization and spray systems*, Vail, CO: ICLASS; 2009.
81. Karami R, Ashgriz N. Flashing sprays. In: Ashgriz N, editor. *Handbook of atomization and sprays*. New York, NY: Springer; 2011.
82. Nandiyanto ABD, Okuyama K. Progress in developing spray-drying methods for the production of controlled morphology particles: from the nanometer to submicrometer size ranges. *Adv Powder Technol* 2011;**22**:1–19.
83. Cal K, Sollohub K. Spray drying technique. I: hardware and process parameters. *J Pharm Sci* 2010;**99**:575–86.
84. Gaspar F, Vicente J, Neves F, Authelin J-R. Spray drying: scale-up and manufacturing. In: Shah N, Sandhu H, Choi DS, Chokshi H, Malick W, editors. *Amorphous solid dispersions*. New York, NY: Springer; 2014.
85. Khavkin YI. *Theory and practice of swirl atomizers*. New York, NY: Taylor & Francis; 2003.
86. Dobry DE, Settell DM, Baumann JM, Ray RJ, Graham LJ, Beyerinck RA. A model-based methodology for spray-drying process development. *J Pharm Innov* 2009;**4**:133–42.
87. Hsieh DS, Yue H, Nicholson SJ, Roberts D, Schild R, Gamble JF, et al. The secondary drying and the fate of organic solvents for spray dried dispersion drug product. *Pharm Res* 2015;**32**:1804–16.
88. Baldinger A, Clerdent L, Rantanen J, Yang M, Grohganz H. Quality by design approach in the optimization of the spray-drying process. *Pharm Dev Technol* 2012;**17**:389–97.
89. Elversson J, Millqvist-Fureby A, Alderborn G, Elofsson U. Droplet and particle size relationship and shell thickness of inhalable lactose particles during spray drying. *J Pharm Sci* 2003;**92**:900–10.
90. Vicente J, Pinto J, Menezes J, Gaspar F. Fundamental analysis of particle formation in spray drying. *Powder Technol* 2013;**247**:1–7.
91. Curatolo W, Nightingale JA, Herbig SM. Utility of hydroxypropylmethylcellulose acetate succinate (HPMCAS) for initiation and maintenance of drug supersaturation in the GI milieu. *Pharm Res* 2009;**26**:1419–31.
92. Chen J, Ormes JD, Higgins JD, Taylor LS. Impact of surfactants on the crystallization of aqueous suspensions of celecoxib amorphous solid dispersion spray dried particles. *Mol Pharm* 2015;**12**:533–41.
93. Dhumal RS, Shimpi SL, Paradkar AR. Development of spray-dried co-precipitate of amorphous celecoxib containing storage and compression stabilizers. *Acta Pharm* 2007;**57**:287–300.
94. Zhang SW, Yu L, Huang J, Hussain MA, Derdour L, Qian F, et al. A method to evaluate the effect of contact with excipients on the surface crystallization of amorphous drugs. *AAPS PharmSciTech* 2014;**15**:1516–26.

CHAPTER

31

Process Development and Scale-Up: Twin-Screw Extrusion

B. Chen[1], *L. Zhu*[1,3], *F. Zhang*[2] *and Y. Qiu*[1]

[1]AbbVie, Inc., North Chicago, IL, United States [2]University of Texas, Austin, TX, United States
[3]Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, PR China

31.1 INTRODUCTION

Extrusion processing is intended to physicochemically transform powder into molten or wet mass and enable the continuous production of highly homogeneous products through the accurate control of processing conditions. It is used to produce a wide range of food and feed products, ceramic, metal, plastics, chemicals, and more recently pharmaceutical products. There are different types of extrusion machines depending on the material extruded or the extrusion technology. Extrusion equipment used today to process pharmaceutical products typically includes extruders with a single or twin-screw auger powered by an electric motor (Fig. 31.1). While single-screw extrusion (SSE) is a mechanically simple and low-cost processing option, twin-screw extrusion (TSE) has been used extensively for mixing, compounding, or agglomerating a variety of ingredients to produce a wider range of products.

In manufacturing solid dosage forms, TSE can be used for hot-melt extrusion (HME), wet and melt granulation. During HME and wet or melt granulation, liquids and powders can be fed accurately and continuously throughout the run to ensure consistency of the formulation. HME is one of the most common processing options to enable the formation of amorphous solid dispersion (ASD) for enhancing dissolution and bioavailability of poorly water-soluble drugs. It is a solvent-free process that transforms a powder blend of active pharmaceutical ingredients (APIs), polymers, and other excipients into an extrudate containing a molecularly dispersed API in a polymer matrix through either melting and mixing the molten drug with molten polymer or dissolving the API in the molten polymer(s) during the extrusion process. The extrudate can be subsequently directly shaped or converted into granules or pellets for downstream processing. Wet and melt granulation via TSE refers to the process of mixing powdered excipients and APIs together with a liquid or meltable binder in the extruder to produce granules with characteristics suitable for manufacturing solid dosage forms. It is more scalable and capable of processing materials consistently with greater process flexibility and higher efficiency when compared to the traditional low-shear, high-shear, fluid-bed, or spray-congealing process. Due to the high mixing efficiency and accurate process control, extrusion granulation usually requires less quantity of binder than other processes. It can be optimized to generate more uniform granules with respect to physicochemical characteristics. It also offers a great flexibility in producing granules with diverse characteristics and requires significantly less materials for process development, scale-up, and optimization. As shown in Table 31.1, the TSE process has been successfully applied to manufacture a growing number of commercial drug products covering a wide range of therapeutic areas.

In essence, TSE offers many advantages over conventional pharmaceutical process, including: (1) continuous processing with increased efficiency and productivity and the ability to handle multiple processing functions in series (such as conveying, melting, distributive and dispersive mixing, granulating, and devolatilizaton); (2) solvent-free processing with a small equipment

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00031-5

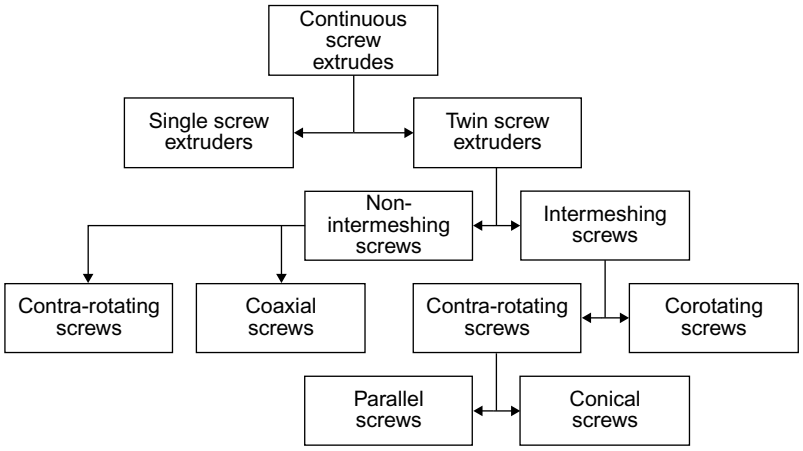


FIGURE 31.1 Types of extruders. *Tangram Technology Ltd.*

footprint; (3) considerable flexibility in achieving specific mixing characteristics to meet different product and process requirements and produce a wide range of pharmaceutical dosage forms with fewer unit operations; and (4) exceptional mixing capability for achieving superior product homogeneity and content uniformity. When the extrusion process is used as a continuous processing technology in commercial production, it is essential to establish and implement a sound control strategy, such as real-time release testing (RTRT). However, in pharmaceutical manufacturing, a process train typically consists of a combination of batch (eg, blending, coating) and continuous unit operations (eg, compression). Thus, batch and lot definition becomes a key consideration. In most cases, a batch can be defined based on the limiting factor(s) of the upstream or downstream unit operations. For a process train involving TSE, batch size of the preblend can often be used to determine product batch size, enabling thus converting back-to-back extrusion runs into a semicontinuous process to meet the requirement for uniform quality while preserving the advantage of high manufacturing efficiency.

This chapter describes the design and basic principles of (1) the TSE process, (2) the design and functionality of the extruder and screw elements, (3) process parameters, and (4) material considerations in process development. The fundamentals of product and process development using TSE are presented, including the HME process for ASD products, and continuous wet and melt granulation for immediate-release (IR) and extended-release (ER) dosage forms. Process scale-up and critical issues are also discussed, including general principles, common approaches, and modeling. Case studies are provided to illustrate the versatility of TSE and some of its successful applications in manufacturing different types of solid dosage forms.

31.2 TWIN-SCREW EXTRUDER AND EXTRUSION PROCESS

31.2.1 Extruder design and components

An extruder can be roughly divided into three sections based on functionality: (1) the drive section, which includes the motor, gear box, shaft linkage, and thrust bearings; (2) the process section, which is composed of the barrel, screw, optionally an endplate/die, and various sensors; and (3) the control section (ie, human machine interface, HMI). The drive and control sections of the extruder are generally defined by equipment specifications, and modification of these portions for individual process requirements is not common. This chapter will mainly focus on the process section of the extruder.

There are two common types of commercially available extruders: the ram type and screw type. Compared to a ram extruder, a screw extruder is preferred due to its continuous manufacturing capability. Depending on the number of screws, screw extruders can be classified as SSEs, TSEs, and multiscrew extruders. SSE is a relatively simple process, and it has served as a fundamental operation for polymer processing since the early 1900s. Due to its superior mixing capability and shorter material residence time, a TSE is more preferable for the pharmaceutical industry. A TSE consists of two parallel rotating screws inside a stationary cylindrical barrel. The two screws can rotate either in the same direction (ie, corotating) or in opposite directions (ie, counterrotating). When a counterrotating extruder is employed, materials are subject to high-shear regions; thus, this setup should be considered when higher shear regions are required. Based on the centerline distance between the two screw shafts, TSEs can be

TABLE 31.1 Commercial Drug Products Manufacturing Using TSE

Product	Indication	Type of dosage form	Company	Year approved
Lacrisert (HPC rod)	Dry eye syndrome	Shaped system Ophthalmic insert	Merck	1981
Zoladex (goserelin acetate implant)	Prostate cancer	Shaped system Subcutaneous implant	AstraZeneca	1996
NuvaRing (etonogestrel, ethinyl estradiol)	Contraceptive	Shaped system Vaginal ring	Merck	2001
Implanon (etonogestrel)	Contraceptive	Shaped system Vaginal ring	Merck	2006
Ozurdex (dexamethasone)	Macular edema	Shaped system Ophthalmic implant	Allergan	2009
Zithromax (azithromycin)	Antibiotic	Sustained release	Pfizer	1996
Palladone (hydromorphone HCl)	Pain	Sustained release Abuse deterrence	Purdue Pharma	2004
Nucynta (tapentadol)	Pain	Sustained release Abuse deterrence	Depomed Inc.	2011
Opana ER (oxymorphone HCl)	Pain	Controlled release	Endo Pharmaceuticals	2011
Eucreas (vildagliptin, metformin HCl)	Diabetes	Immediate release	Novartis	2007
Gris-PEG (griseofulvin)	Antifungal	Crystalline dispersion	Pedinol Pharmacal	1975
Rezulin (troglitazone)	Diabetes	Amorphous dispersion	Wyeth	1997
Kaletra (ritonavir/lopinavir)	Antiviral (HIV)	Immediate release Amorphous solid dispersion	AbbVie	2005
Norvir (ritonavir)	Antiviral (HIV)	Immediate release Amorphous solid dispersion	AbbVie	2010
Onmel (itraconazole)	Antifungal	Immediate release Amorphous solid dispersion	Merz North American, Inc.	2010
Noxafil (posaconazole)	Antifungal	Delayed release, Amorphous solid dispersion	Merck	2013
Viekira Pak (ombitasvir, paritaprevir, ritonavir, dasabuvir)	Antiviral (HCV)	Immediate release Amorphous solid dispersion	AbbVie	2014
Belsomra (suvorexant)	Insomnia treatment	Immediate release Amorphous solid dispersion	Merck	2014
Zepatier (elbasvir and grazoprevir)	Antiviral (HCV)	Immediate release Amorphous solid dispersion	Merck	2016
Venclexta (venetoclax)	Oncology	Immediate release Amorphous solid dispersion	AbbVie	2016

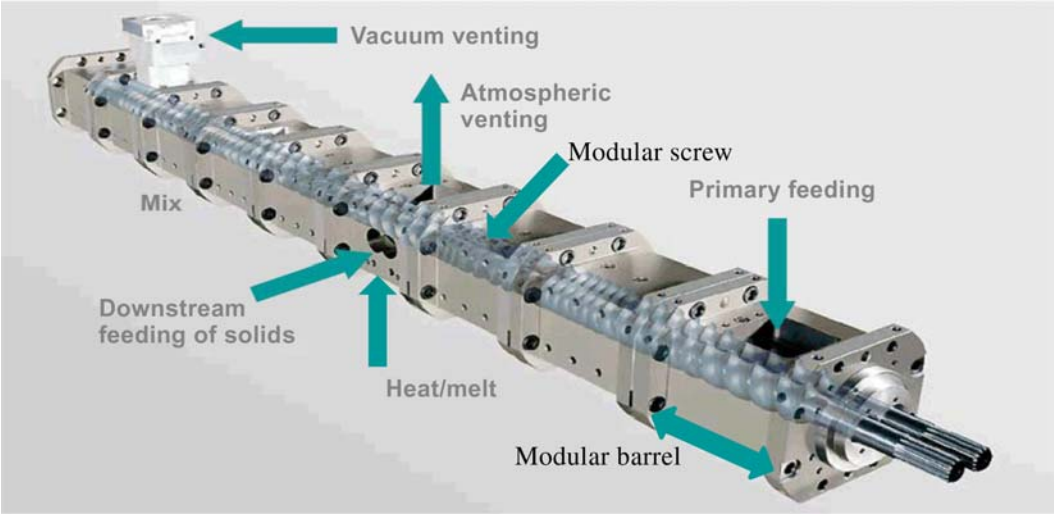


FIGURE 31.2 Extruder barrels and screws.[1]

further classified as fully intermeshing TSEs (ie, the centerline distance is less than the screw diameter) or nonintermeshing TSEs (ie, the centerline distance between two screw shafts is equal to the screw diameter). The detailed classification of screw extruders can be found in Fig. 31.1. Fully intermeshing corotating TSE offers self-wiping or self-cleaning feature, and is the most widely used and studied setup for pharmaceutical applications.

31.2.1.1 Barrels

Extruder barrels are the housing for the screws. The barrels are generally electrically heated and water cooled. During the extrusion process, the barrel temperature is monitored by thermocouples. Depending on the design and manufacturer, the barrel can be either a one-piece or a modular design. Compared to the one-piece design, modular barrels are more widely used in the pharmaceutical industry due to their greater flexibility to meet processing or product requirements. Modular-designed barrels and screws are shown in Fig. 31.2. The typical length of an extruder should be determined based on the screw configuration required for the process. However, the individual modular barrel can be in different lengths and diameters. The barrel temperature can also be controlled individually or in grouped zones. In addition, the modular barrels could be configured differently (see Fig. 31.3).[1] Some common barrel configurations are closed barrel, feed barrel, vented barrel, and combi barrel.[1] The closed barrels generally comprise the majority of an extruder design. Liquid feed, if needed, can be introduced into the system through an injection nozzle connected to a closed barrel or a combi barrel.

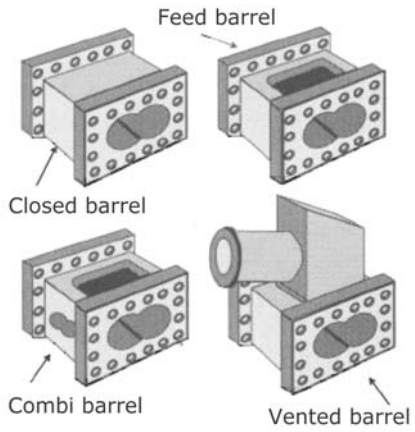


FIGURE 31.3 Different types of barrel section.[1]

The feed barrel, with either a circular- or rectangular- shaped opening on the top, is generally the first section in the extruder and is normally kept near room temperature. To maximize the free volume for material to enter the extruder, the screw in the feeding section should have a long pitch to keep the material moving forward rapidly (see detailed discussion on screw design in Section 31.2.2). However, depending on formulation, the number and locations of the feed barrels should be carefully considered. Adding material downstream reduces its residence time and the amount of shear the material experiences. Bruce et al. reported that using a side-stuffing port close to the die could successfully incorporate heat-sensitive

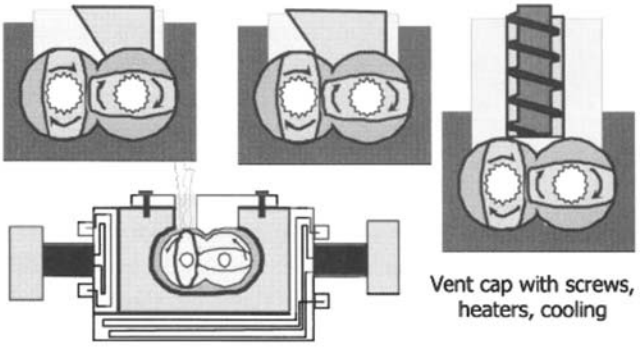


FIGURE 31.4 Vent cap configurations.[1]

dextromethorphan hydrobromide granules with a molten polymer plasticizer while avoiding exposure of the material to the high process temperature required for melting the carrier polymer.[2]

The vented barrel is used to remove air and/or volatiles from the extruder; this process is also referred as devolatilization. The vent barrel also possesses an opening on the top, which can be connected to atmosphere or vacuum. Devolatilization may be critical to the extrusion process. The proper venting (or devolatilization) should allow air and/or volatiles to escape from the barrels while preventing the material from being lost. A variety of vent-port designs, including vent stuffers, elongated vent barrels, and side vents, have been developed for corotating TSEs to accommodate a wide range of applications and materials. In general, the vacuum port covers approximately half of the screw surface; this can be achieved by either the vent port design or addition of vent caps. Some typical vent cap configurations are shown in Fig. 31.4.[1] Besides regular-shaped openings, the barrel can also house dynisco ports (ie, standard high-pressure/temperature instrumentation ports) for installing various process monitoring probes.

31.2.1.2 Screw elements

Screw design can significantly impact product characteristics and quality because it affects mechanical shear and residence time.[3] Thompson and Sun reported that changes in screw elements could influence the size and shape of lactose-polyvinylpyrrolidone granules for a wet granulation process.[4] Nakamichi et al. showed that a kneading paddle play an important role in the preparation of ASD of nifedipine.[5] Therefore, special attention should be given when designing screws. The design must consider both the process requirements and the material properties. Screw configurations can be designed as shear-intensive or shear-passive according to the process requirements. With modular designs (see

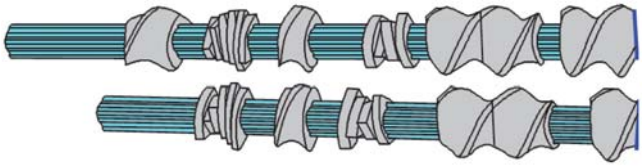


FIGURE 31.5 Modular screw design.[118]

Fig. 31.5),[118] there are seemingly infinite possibilities of screw variations. However, considering the discrete zones, including the solid feed/conveying zone, mixing/melting zone, the melt conveying/metering zone, and potential devolatilization/venting zones as shown in Fig. 31.6 for a basic extrusion operation, there are predominately two types of screws: (1) flighted (threaded) screws to achieve conveying and (2) non-flighted screws for mixing and melting by creating shear regions.[6,7] One example of a complete modular screw design is shown in Fig. 31.6.[8] During extrusion, material can be transported mainly through drag flow and pressure flow. The pressure flow is due to the pressure buildup at specific regions, while the drag flow is created from the volumetric displacement of screw.

31.2.1.2.1 Conveying screws

One typical flighted screw is the conveying screw. The helix geometry of a single-flight conveying screw is shown in Fig. 31.7.[9,10] The pitch is defined as the distance between consecutive flights. The helix angle is the angle between the screw flight and the plane perpendicular to the screw axis. The pitch is directly related to the helix angle. The optimum helix angle/pitch should be designed based on the material properties and design requirement. Fight clearance, with a rule of thumb of 0.1% of the screw diameter, is the space between the screw outer diameter D_o and the barrel wall. D_i is the inner screw diameter.

Conveying screws are the primary constituent of the screw configuration in the feed zone and metering zone. The feed section is responsible for conveying materials such as pellets and powder forward to the transition section. The metering section, also called melt conveying or pumping zone, delivers the melt toward the discharge end of the extruder. The helix geometry enables material transportation through the open channel downstream. The capability of the screw conveying is directly related to the volumetric displacement or free volume of the channels. The screw conveying capacity can be calculated in many different ways. One example is shown in Eq. (31.1); the conveying capacity is expressed as drag flow Q_d during the extrusion process.[10]

$$Q_d = 0.5 \times a \times z \times \cos(\theta) \times N \tag{31.1}$$

FIGURE 31.6 Example of a complete modular screw design.[6,118]

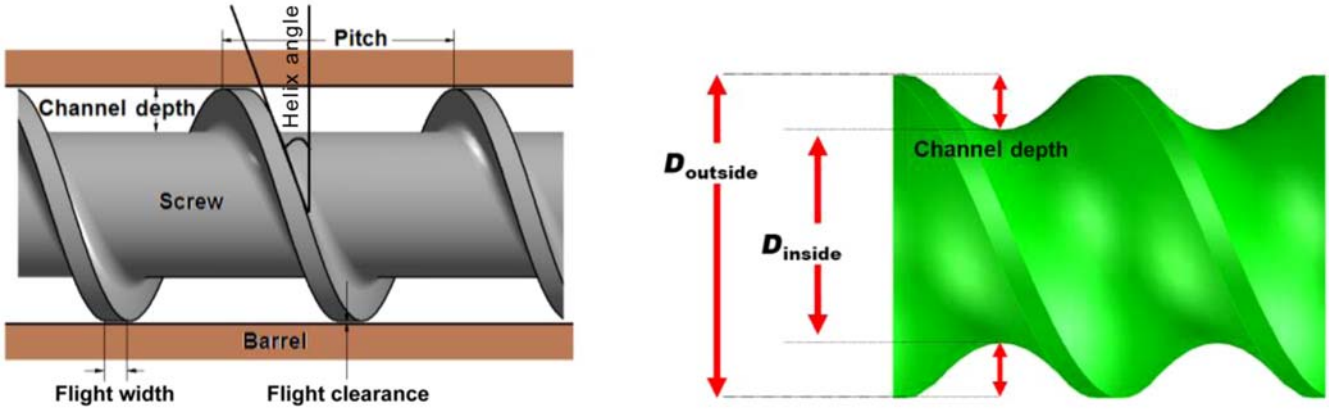


FIGURE 31.7 Basic geometry of a conveying screw element.

where Q_d is the drag flow; α is the area of the open section of the screw; z is the lead length of the screw; θ is the helix angle; and N is the screw speed. The lead length is proportional to the screw pitch. For a single-flight screw element, the lead length is equal to pitch (ie, the axial distance that screw travels in one complete revolution of 360°), while for a double-flighted screw, the lead length would be twice of the pitch.

Regardless of the various mathematical expressions, it is clear that the screw conveying capacity is primarily a function of the screw open section or the free volume. The screw design, including the number of flights, ratio between D_i and D_o, and pitch, has a significant impact on the free volume. Zimmermann listed the free volume of various screw designs.[11] The comparison among three commonly used conveying screw designs is shown in Table 31.2.[118] The impact of pitch on volumetric displacement and general guidance of its use are summarized in Table 31.3.[118] The screws in the feeding zone usually have a greater pitch; the pitch reduces in the metering/compression zone to increase pressure by increasing the degree of fill while removing entrapped air. For a HME process, the screw with a pitch less than one screw outer diameter ($<D_o$) is employed in the melt-conveying section to generate pressure and pump the melt uniformly out to the die cavity. As the degree of fill could also impact the average shear rate (energy input), as shown in Fig. 31.8,[12] the optimum design should be defined based on the main function of the processing zone and rheology of the materials being processed.

TABLE 31.2 Comparison Among Common Conveying Screw Design[118]

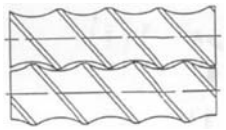
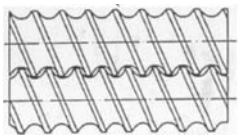
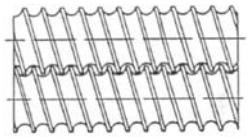
Screw design	Free volume	Conveying efficiency	Pros	Cons
	+	+++	Low leakage flow High pumping efficiency	Lower conveying capacity
	++	++	Good balance between free volume and conveying efficiency	Higher shear than one-flighted screw
	+++	+	High conveying capacity	Some back-mixing

When evaluating the screw design, the compression ratio is an important parameter that should be assessed. The compression ratio is the ratio of the channel depth (or channel volume) in the feeding zone to that in the metering zone. This ratio is closely related to the slope of the transition section. During HME, the compression ratio and the close of the transition section should be carefully matched with the polymer melting rate to ensure an efficient process. A screw with a high compression ratio is generally recommended for processing materials with high compressibility.

31.2.1.2.2 Mixing screws

Although typical helix conveying screws provide efficient material transportation, they may not be able to provide sufficient distributive and dispersive mixing, which is particularly important for generating a molecularly dispersed formulation. Distributive mixing is a low-shear process; it promotes homogeneity by diving and recombining the material without significantly reducing particle size. In contrast, dispersive mixing is a high-shear process; it breaks up large particles and disperses them throughout the material. Mixing that accompanies melting is dispersive in nature, as the high-shear stresses open an opportunity for the elongation and breakup of material. A diagram of both distributive and dispersive mixing is shown in Fig. 31.9.[13] Staggered kneading blocks are widely employed mixing element configurations. A kneading block element generally consists of several kneading disks. The number and width of kneading disks, as well as the stagger angle (offset angle) between successive disks, are the three most important

TABLE 31.3 Impact of Conveying Screw Pitch[118]

Screw pitch	Volumetric displacement	Degree fill	Uses
High (1.5–2 D_o)			
	+++	+	Feeding materials into extruder
Medium (1–1.5 D_o)			
	++	++	Compression after Feeding Melt transport
Low (<1 D_o)			
	+	+++	Pressure generation Melt pumping

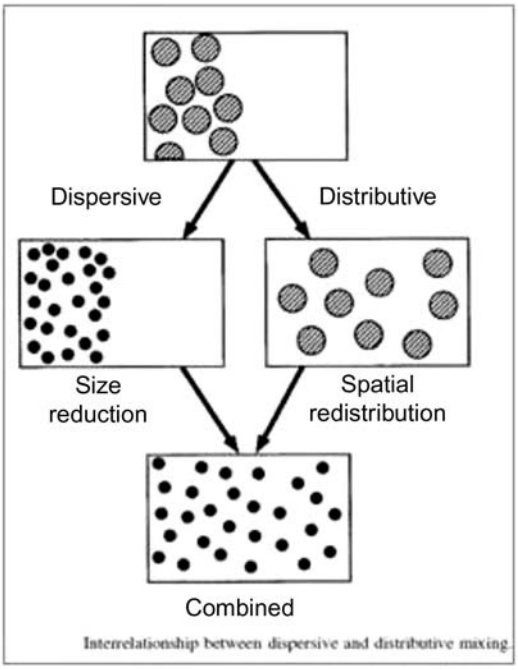
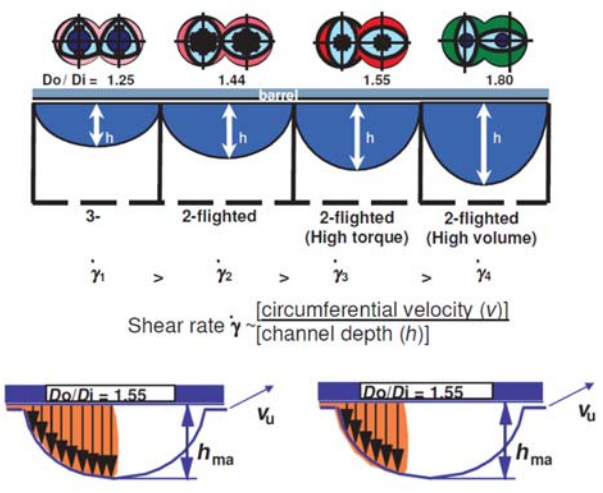


FIGURE 31.9 Comparison of distributive and dispersive mixing.[13]



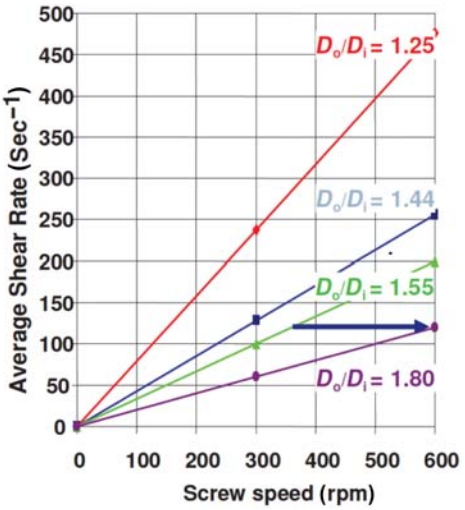


FIGURE 31.8 Impact of D_o/D_i on shear rate.[12]

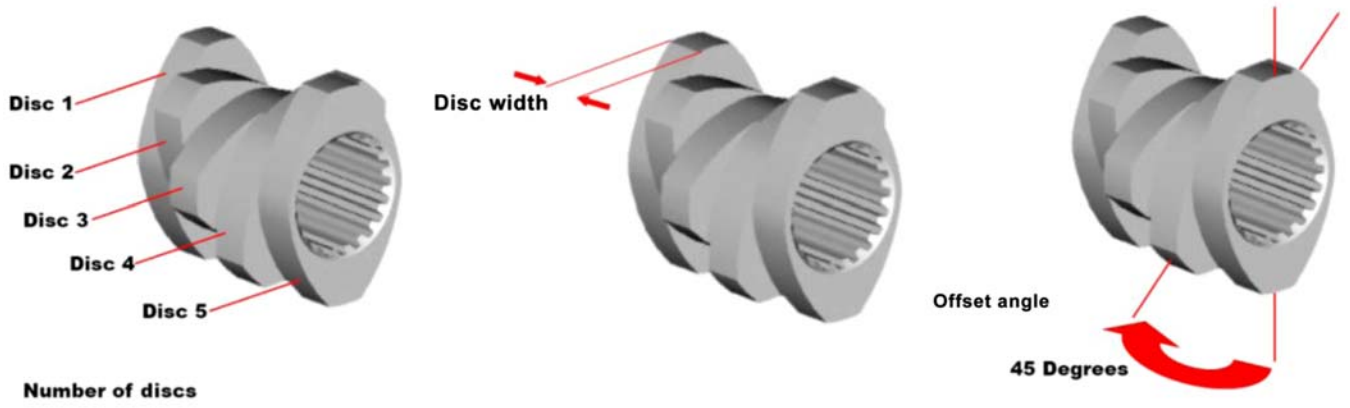


FIGURE 31.10 Geometry of a kneading block.[118]

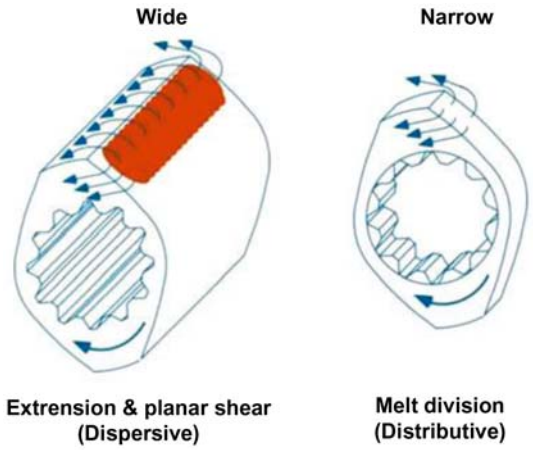


FIGURE 31.11 Material flowing around and over a kneading block.[14]

characteristics for kneading blocks. The geometry of kneading block elements is presented in Fig. 31.10. Depending on the design, kneading blocks create various shear intensities, and they can be used for either distributive or dispersive mixing. Narrow kneading disks provides good distributive mixing as more material is forced to flow in and around blocks instead of between the block and barrel wall. Wider kneading disks are mainly used for dispersive mixing because material tends to go over the top of the discs. Material flow around and over kneading disks is illustrated in Fig. 31.11.[14]

The kneading block could be configured as forward (or right-handed), neutral, and reverse (or left-handed) kneading elements (as shown in Fig. 31.12) based on the stagger angle. Since the stagger angle between disks causes an axial displacement flow, both forward- and reverse-kneading blocks also have conveying capability. However, this type of conveying, as it is related to the stagger pressure profile, is different from the drag flow observed with conveying screw elements. The conveying direction of the forward kneading elements is the same as the direction of the product flow in the extruder. It allows back-mixing of material through the gaps formed by the staggering angle. The conveying direction of the reverse kneading elements is contrary to the direction of the product flow in the extruder. Thus, the aim of incorporating reverse element in screw configuration is to produce back pressure and increase upstream degree of fill. Neutral kneading blocks have no conveying capability. Table 31.4 summarizes a simplified characteristic comparison among different kneading block elements.

In addition to kneading blocks, many other special mixing elements could be considered for improving mixing. A few examples are illustrated in Fig. 31.13.[15] Special distributive mixing elements (DMEs) utilize interrupted screw flights to allow higher material exchange between adjacent screw channels, thus enhancing back-mixing. As the opening on the screw flights causes increased leakage flow, the conveying efficiency of the screw is reduced. However, a higher fill factor and longer residence time are expected. Mixing gears or toothed mixing elements (eg, Tooth Block, ZME) offer the maximum amount of distributive mixing with minimal energy input by creating flow splitting and reorientation. Mixing gears consist of a series of cams, which divide the material flow and form multiple layers by the swirling action. The interfacial area is thereby increased. The material layers are then wiped between the cam tips and barrel, resulting in enhanced distributive mixing. The mixing efficiency of mixing gears is impacted by the number of gears and cams per gear. The degree of homogenization increases with increasing the number of mixing gears and/or the number of cams per gear. Barrier screws

FIGURE 31.12 Forward and reverse kneading blocks.

TABLE 31.4 Characteristics of Kneading Blocks[118]

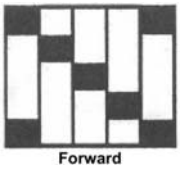
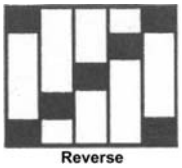
	 Forward	 Reverse	 Neutral	 Forward	 Reverse	 Neutral
Conveying efficiency	+++	−−−	None	+	−	None
Distributive mixing	+++	++++	++++	+	++	++
Dispersive mixing	+	++	+++	++++	++++	+++++
Shear input	+	++	++	+++	++++	+++

FIGURE 31.13 Examples of special mixing elements.[119]

are believed to provide gentle dispersive mixing by passing particles through a defined barrier gap and exposing them to a low and well-defined shear stress. For HME, the compression and decompression of the material lead to a complex elongation flow and, therefore, lead to an optimal dispersive mixing effect. Kessler et al. demonstrated the suitability of applying barrier screw elements for manufacturing ASD by promoting mixing at relatively low temperature.[16]

The ratio between screw length and screw diameters (L/D) is an important characteristic to consider, as it impacts the material residence time and the length assigned to each process operations. This ratio, to a large extent, is determined by the extruder design. The typical L/D range for intermeshing TSEs is from 20:1 to 40:1, with a potential for up to 60:1. Higher L/D ratios offer more flexibility in terms of screw design to meet specific process requirements; however, it should balance with the consideration of residence time and torque. For pharmaceutical applications, the pilot scale screws have diameters ranging between 18 to 30 mm, whereas the screw diameter for production scale could exceed 50 mm. In advanced extrusion equipment, the screw dimensions can change over the length of the barrel and extend proportional to the length of barrel.[17–20] Morott et al. (2015) demonstrated the impact of L/D ratio along the screw design on taste-masking of sildenafil citrate.[21] A shorter screw design

with one mixing zone enhanced the perseveration of the API crystalline nature.

31.2.1.3 Dies

The main function of an extrusion die is to shape the pressurized molten material to the desired shape and desired dimensions by passing it though a predefined cross-section.[1,22,23] The die is attached to the end of the extruder and provides a passage for the material exiting from the circular extruder barrel to a more complex and often much thinner and wider die exit. A die is generally not required for extrusion granulation. Extrusion dies may vary dramatically in shape and complexity to meet the demands of the product being manufactured. For melt extrusion, an ideal die should be able to (1) balance melt flow to provide acceptable die pressure drop and (2) keep melt homogeneity without causing stagnated flow. The stagnation area could lead to thermal degradation, as the melt is being exposed to high heat for an extended period of time. Strict geometric shaping may not be required if the extrudate will be milled into fine powder for further downstream processing. However, for melt extrusion, the material could spend a significant portion of its residence time in the die under thermal stress, thus, the proper design of an extrusion die is important and requires the understanding of material characteristics, knowledge of mass and heat transfer, and extensive experience with extrusion processing. A die with a large open cross-section is believed to minimize die pressure while maintaining appropriate melt linear velocity and is thus recommended for general use.

Materials' rheological properties should be carefully evaluated when designing a melt-extrusion die, as these properties could significantly impact product quality. For example, die swelling (also known as extrudate swelling or the Barus effect), as a result of velocity relaxation of melt flow and viscoelastic relaxation of strained polymer molecules, could cause a change in the melt flow velocity profile exiting from the die and result in a stagnant region.[23–25] Giles described the four zones of the extruder die.[1] The parallel zone, where the melt acquires the final characteristics and shape prior to exiting from the die, is believed to mainly control the die swell, back pressure, and flow uniformity. To compensate for material and process variability, appropriate die adjustments, such as mandrels, choker bars, die lips, and localized temperature control, could be required to maintain a robust and stable process. Depending on the final application, slit dies (for film and sheet), strand dies (for simple geometric shapes, such as circles, squares, or triangles), and annular dies (for pipe and tubing) are commonly used dies in the pharmaceutical industry. Historically, extrusion die design is more an art than a science. However, it has been improved with recent advancements in powerful computation tools such as finite element analysis (FEA) for modeling of complex flow and heat transfer processes.

31.2.1.4 Auxiliary systems

A wide range of auxiliary systems, including feeders and various downstream equipment, are generally used with the extruder to complete the extrusion process. The selection of feeder and downstream processing equipment depends upon many factors such as desired dosage form, material characters, and processing speed. Feeders with precise and controllable discharge rates are employed to accurately meter materials into the extruder. Preblend materials can be added via a single hopper or, alternatively, multiple feeders can simultaneously add different ingredients in the correct ratios. The feed rate can be controlled either by volume (volumetric feeder) or by weight (gravimetric feeder). Volumetric feeders are designed based on some form of displacement principle, such as a rotary feeder or rotating screw. It is an open-loop device and cannot directly measure material throughput to compensate for dosing variation. The dosing accuracy of volumetric feeders is within 1–3%. In contrast, gravimetric feeders control material feed by weight and improve the dosing accuracy to 0.5–1% by adding a weigh system and closed-loop control scheme. Based on the actual discharging weight over time, the system makes necessary feeder speed adjustments to maintain the desired mass feed rate. Gravimetric feeding is the most preferable for pharmaceutical applications primarily due to better dosing accuracy. However, material properties, such as density, flowability, particle size, frication, and compressibility, play a critical role in feeder selection. For example, a crammer feeder may be a viable option when dealing with materials that have a low bulk density. If necessary, liquid additives can also be introduced to a TSE through an injection port with a typical a gravimetric or loss-in-weight liquid feed pump.

Downstream auxiliary equipment is utilized to achieve the desired dosage form. There are a wide variety of downstream systems available to extrude an infinite array of shapes. Fig. 31.14 presents different dosages that could be produced by an extrusion process.[26]

For melt extrusion, direct shaping can be achieved by either extruding melt direct to a mold with specific dimension by calendaring.[27] During calendaring, the molten strand leaving the extruder is shaped to the desired dimensions between two counterrotating calendar rollers. If applicable, direct shaping could be a cost-effective route for dosage manufacturing, mainly due to its continuous manufacturing nature. The suitability of the direct-shaping process could be highly

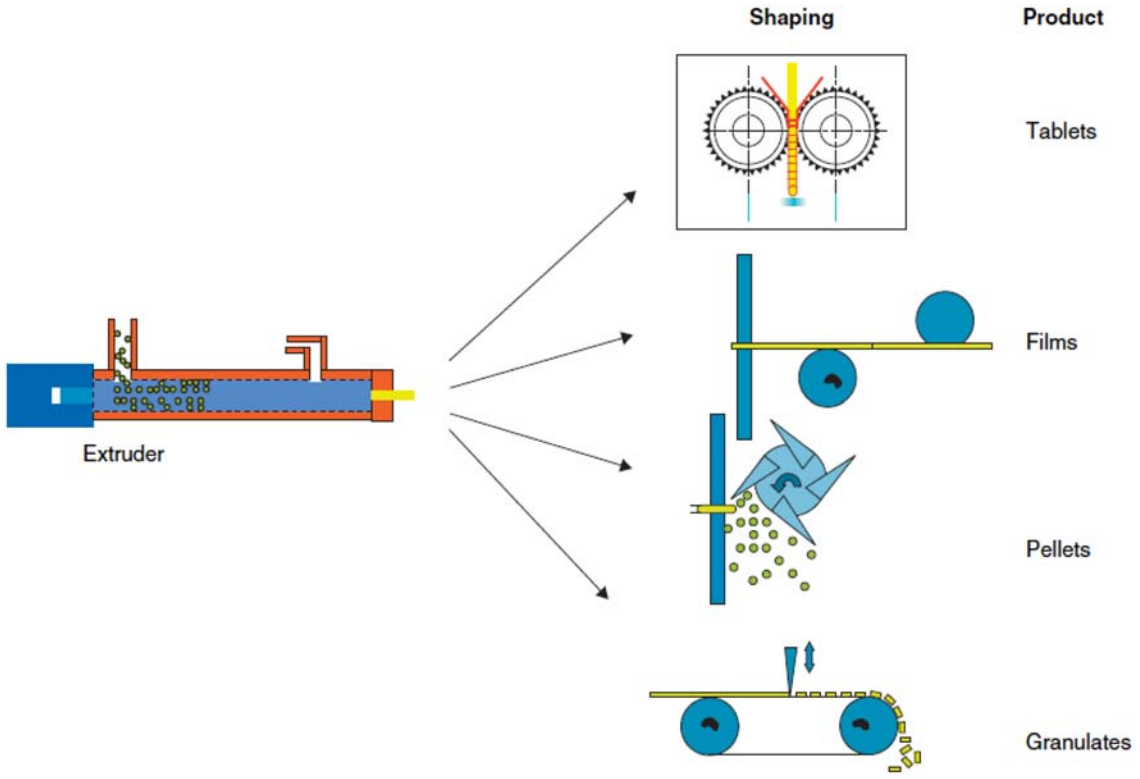


FIGURE 31.14 Different extrudate shapes produced using a twin-screw extruder.[26]

formulation dependent. One example of direct shaping is Egalet, a novel erosion based controlled-release delivery system, which is manufactured via injection molding.[28,29] Roth et al. also reported directly shaping an abuse-resistant, melt-extruded formulation into elongated tablets.[30] They believed that direct shaping increased the tablet strength by eliminating the sharp features that result from the extrudate grinding step. However, DiNunzio et al. claimed that direct-shaped dosages could suffer from a reduced erosion rate due to the low surface area and near-zero porosity; thus, it is rarely employed.[31]

As a post extrusion milling operation is often required for extrudate to manufacture the final dosage, tight dimensional control of extrudate is generally not required. Pellets are one of the most commonly used shapes and can be formed through either calendaring or pelletization.[23,32] Producing sheets of pellets by calendaring is straightforward. During calendaring, the melt is forced to distribute between two roller faces. Depending on the design of the roller surface (smooth surface vs surface with cavities), either films or sheets of pellets can be obtained. The shape and size of the pellets are determined by the roller surface cavities. The temperature of the roll surface is maintained by liquid circulation through internal channels. The calendaring conditions could play an important role in determining the physical properties and appearance of the film/sheets. When the calendar temperature is too low, the melt is difficult to stretch; when the calendar temperature is too high, the material could stick to the roller. In addition, the gap between calendaring rollers could also influence the film shape. Pelletization is a process to form extrudate pellets with typical size between 0.5 and 5.0 mm. Strand pelletization and die-face pelletization are two common pelletization methods. They distinguish themselves in terms of material temperature where the cutting takes place.[33,34] In strand pelletization, spaghetti-like melt strands are pulled by a pelletizer feeder while cooled and pushed to a cutting assembly to size to the required length. Thus, the material is generally cut at a temperature below its softening point, and pellets are more cylindrical-shaped. The diameter of the final pellet is controlled by the diameter of die orifice and pulling speed of the feeder roller, with a typical size of ≤ 3 mm for pharmaceutical applications. Die-face pelletization forms the pellets by direct cutting the melt at the die face and then cooled, most likely by air. Spherical pellets can be obtained through this process due to surface tension related deformation. Die-face pelletization offers major advantages in terms of process simplicity and spatial flexibility; therefore, if amenable, it is often preferred over the

strand system. However, since the material is cut while in its viscous state, this process could be limited in range of applicable formulations mainly due to stickiness.[35]

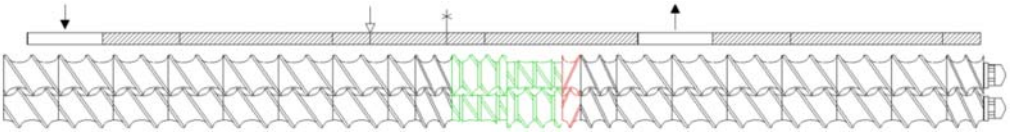
31.2.2 Extrusion process design

Extrusion is an integrated unit operation composed of several function zones. Process parameters play a key role in determining the properties of the extruded product. The complexity of the material and energy properties in different functional zones throughout the extruder is illustrated in Table 31.5. To clearly understand the overall impact of different types of parameters associated with the process on product quality attributes, analytical descriptions of the extrusion unit operation have been reported in the literature as shown in Fig. 31.15.[36] Based on this analysis, the product quality is controlled directly by key system parameters including specific energy, melt temperature, residence time, and pressure, while the common adjustable equipment and process parameters impact the product quality indirectly. In this section, the key process considerations based on the adjustable parameters (eg, feed rate, screw speed, temperature profiles) will be discussed to provide an overview of how the process variables may affect the properties of extrudate.

31.2.2.1 Feed rate

Feed rate is one of the most important adjustable process parameters. During operation, the extruder can be force-fed or starve-fed. When operated under force-fed mode, material is positively displaced into the extruder. When operated under starved-fed mode, the extruder is fed at a rate below the conveying

TABLE 31.5 Material and Energy Properties Throughout Extruder



Function zone	Solid feed	Solid convey	Melting	Mixing	Melt convey	Discharge
Material sate	Solid	Solid	Solid/Melt	Melt	Melt	Melt
Pressure (Bar)	1	1	1	>1	>>1	>>1
Conductive energy	Out of material	In and out of material	Out of material	Out of material	Out of material	Out of material
Energy conversion	None	None	PED/FED	VED	VED	VED

PED, plastic energy dissipation; FED, frictional energy dissipation; VED, viscous energy dissipation.

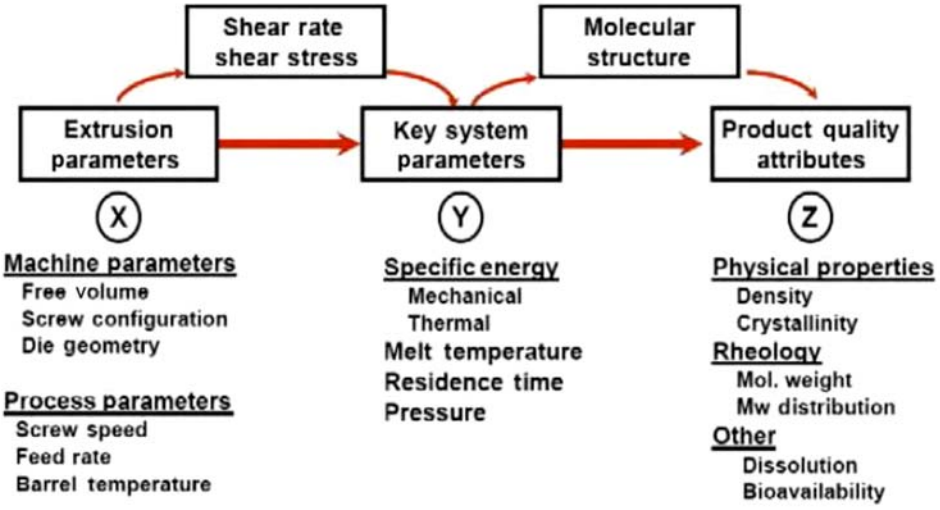


FIGURE 31.15 Analytical description of extrusion unit operation.[36]

capacity of the extruder, and thus a few of the functional zones will not be completely filled. In a starved-fed machine, the mass flow rate at the feed zone is equal to the mass exiting the barrel and accumulation in the barrel is negligible; thus, extrusion throughput is mainly determined by the feed rate but not the screw speed.

TSEs are normally run starve-fed to improve mixing and melting efficiency. The feed rate impacts the extrudate quality by affecting residence time and/or degree of fill during extrusion. Residence time represents the time that materials remain in the extruder. A short residence time leads to inhomogeneity, while long residence time can impact the degradation profile of thermally labile compounds. For melt extrusion, increasing the feed rate with an equivalent increase in the die opening generally reduces the material residence time in an extruder, which results in lower specific energy input as well as higher melt viscosity and extrudate density.[37,38] Gao et al. (2012) proved this claim experimentally.[39] Degree of fill is defined as the ratio of the volume occupied by the material to the free channel volume. An approximation of the degree of fill for conveying screw is given in Eq. (31.2).[10]

$$f = \frac{Q}{Q_d} \tag{31.2}$$

where f is the degree of fill, Q is feed rate, and Q_d is drag flow. It is clear that, under constant throughput, the degree of fill is inversely proportional to screw conveying capacity. Since reverse elements convey material in the opposite direction of material flow in the extruder, the degree of fill increases to approximately 100% immediately upstream of a reverse-pitch element due to limited drag flow.[40] This results in a buildup of localized pressure and facilitates melting; the length of this section (ie, melt section or restriction section) is insensitive to feed rate. Thus, increasing the feed rate does not linearly increase the total melt length; a decrease in specific energy input to material is then expected. However, when using a screw configuration composed of a large portion of forward elements, an increased feed rate could cause choking at the small pitch screws. Under such situations, the specific energy input to the material can increase with feed rate. Fig. 31.16 illustrates a conceptual design space of throughput versus energy input to the formulation.[41]

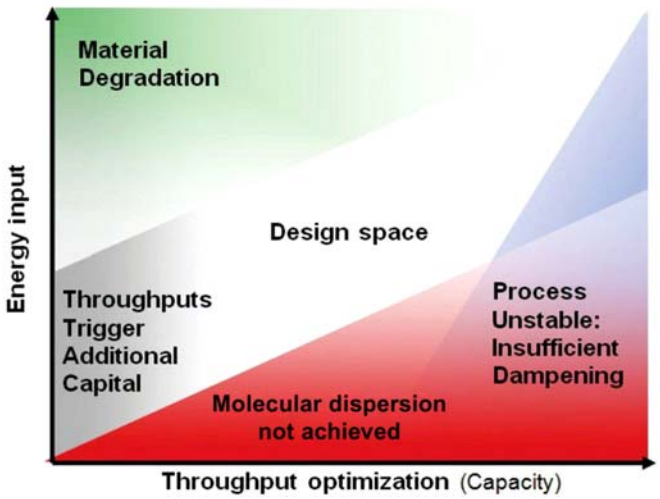


FIGURE 31.16 Conceptual representation of melt extrusion design space.[41]

31.2.2.2 Screw speed

The screw speed in a starve-fed extruder is a critical process parameter for the mixing and melting processes in extrusion. Higher screw speeds increase shear rate, shear stress, and specific mechanical energy (SME) input in all the melt-filled sections. It also reduces the length of the melt pumping section, thus leading to lower motor torque, higher melt temperature, lower die pressure, and higher product expansion.

Shear rate is defined as the velocity gradient between two moving surfaces at different speeds. The peak shear rate can be calculated according to Eq. (31.3).[14]

$$\textit{Peak share rate} = (\pi \times D \times n)/(h \times 60) \tag{31.3}$$

where D is the screw diameter, n is the screw speed in rpm, and h is the overflight gap. When considering mixing of solid particles with molten carriers, shear stress, which is related to the degree of deagglomeration, can then be estimated based on shear rate and viscosity (as in Eq. 31.4).[120] Based on Eq. (31.3), higher screw speeds lead to higher peak shear rates and better deagglomeration/mixing.

$$\textit{Peak shear stress} = \textit{Peak shear rate} \times \textit{Viscosity} \tag{31.4}$$

In addition, shearing of the polymer melt generates viscous energy dissipation. In this process, the mechanical energy from shearing is transformed into thermal energy. The rate of heat generation per unit volume due to viscous heat dissipation follows Eq. (31.5), in which m is a constant, and n is the power law constant.[20]

$$E = m \times \textit{Shear rate}^{n+1} \tag{31.5}$$

Heat in an extruder converted through viscous dissipation can also be expressed as SME. SME describes the mechanical energy input to the extrusion system

per unit mass in the form of work from the motor. Obtaining an accurate SME is difficult due to many uncertainties during measurement. One of the most straightforward calculations is presented in Eq. (31.6).[38,42]

$$SME = \frac{Power}{Throughput} = \frac{P \times \tau \times N}{m \times N_{\max}} \tag{31.6}$$

where P is the motor power or maximum available power in kW. N and N_{max} are the actual screw speed and the maximal screw speed for a giving system, respectively; τ is the torque or motor load expressed as a percentage of the rated (or maximal) power; and m is the real throughput of the process, which equals to the feeding rate in a starved-feeding system. The representative τ of the process should not include the torque level when the extruder is running empty. Assuming the heat capacity of the extrudate is known, the temperature rise can then be calculated using SME. Increasing the heat generated via viscous dissipation favors the crystalline-to-amorphous transition by increasing API dissolution in the matrix polymer and enables the production of hot-melt extruded ASD. However, a high degree of viscous heat dissipation may also induce degradation of drug substances and polymers.

Screw speed has a more significant influence on screw configurations with more restricted elements, as those configurations generally contain a greater portion of reverse paddles/screw elements. Although the degree of fill in restricted and nonforward flow sections does not change with screw speed, the degree of fill with low-conveying-capability screw elements and paddles may rise significantly with screw speed to a certain threshold level.

The impact of screw speed was evaluated on various formulations. Henrist and Remon investigated the manufacturing process for starch-based extrudates.[43] Screw speed and feed rate were found to impact physical and chemical proprieties of the extrudates. The combination of a high screw speed and a high feed rate lead to a large extrudate with low mechanical strength as well as a slow dissolution rate. However, Verhoeven et al. reported that, within the process condition range studied, screw speed had limited impact on the drug release of hot-melt extruded metoprolol tartrate-ethyl cellulose mini-matrices.[44] Shibata et al. suggested that screw speed influences the preparation of solid dispersion of indomethacin−crospovidone mainly by impacting the residence time. However, in the study, they also claimed that the X-ray diffraction (XRD) peak of indomethacin decreased with decreasing screw speed, which is contradictory to the theory that reducing screw rotation speed decreases mechanical shear.[45]

31.2.2.3 *Barrel temperature*

Since heat conduction from the barrel contributes to the melting process, barrel temperature could be critical to control the extrusion process, optimize throughput, and minimize degradation. An optimized temperature profile should be designed based on the understanding of raw material properties and the finish product requirements. Since extrusion is an integrated process, the selection of the temperature profile should also take into consideration of the screw configuration employed and other process parameters chosen.

In general, for melt extrusion, a progressive temperature profile is utilized, that is, the temperature set points increase continually from the feeder to die. In such temperature profiles, the feeding section is water cooled without heating to provide uniform flow without premature melting. The material is preheated and compressed in the solid convey zone. As rule of thumb, the temperature in the solid convey zone should be set 11−17°C below the melt temperature for both semicrystalline and amorphous polymers. The desired melt temperature should be reached at the melting zone and maintained in the remaining barrels. Certain minimum temperatures are required in the HME process to reduce the torque needed to rotate the screw. Typically, the temperature of the melting zone is set 15−60°C above the melting point of semicrystalline polymers or the glass transition temperature of amorphous polymers.[46,47] If additional dispersive and/or distributive mixing is required, the barrel temperature may need to be decreased to remove the excessive heat generated by shear. When devolatilization is needed, a constant temperature during venting is desired to facilitate the removal of volatiles. The temperature of the melt convey and discharge zone should be set to ensure good flowability into the die and acceptable melt strength.

To some extent, barrel temperature manages melt viscosity, which greatly affects the mixing between molten polymer and API. As the Weber number describes in Eq. (31.7), the viscosity of molten carriers plays an important role in the mixing of dispersed phase (ie, molten drug substance) within a continuous phase (ie, molten carrier).[120] Increased melt viscosity favors mixing.

$$W_e = \frac{\gamma \times d_0 \times \eta_c}{2\sigma} \tag{31.7}$$

in which W_e is Weber number, γ represents shear rate, d_0 is the diameter of the molten drug substance before mixing, η_c is the viscosity of the molten carrier, and σ is the interfacial tension between drug substance and carrier.

The impact of barrel temperature on products made by HME has been reported in several studies. Henrist

et al. found that only the highest barrel temperature influenced the characteristics of the starch-based extrudates.[43] Low temperatures caused a rough surface, while high temperatures induced foaming, with a temperature of 100°C being considered as optimum. When extruded at high temperature, Shibata et al. observed the shift of indomethacin absorption peaks in indomethacin-crospovidone solid dispersions, which contributed to the molecular interaction changes.[45] They also claimed that high extrusion temperature could enhance the dissolution rate of the formulation by inducing drug supersaturation.

The melt viscosity is highly dependent on both the glass transition temperature (T_g) of the carrier polymer and the processing temperature, as shown in the WLF Eq. (31.8).[121]

$$log\left(\frac{\eta}{\eta_{T_g}}\right) = \frac{-C_1(T-T_g)}{C_2+(T-T_g)} \tag{31.8}$$

in which η and η_c are the melt viscosity at a processing temperature and the glass transition temperature, respectively; C_1 and C_2 are constants; and T_g is glass transition temperature. Based on the equation, the processing temperature may have to be set relatively high to achieve a desired melt viscosity. However, the high extrusion temperature may cause the degradation of drug substances and/or polymers. Therefore, formulation and process are closely related and must be considered together when selecting a barrel temperature. Verreck et al. explored the possibility of employing supercritical carbon dioxide as a transient plasticizer for itraconazole-copovidone and p-aminosalicylic acid (p-ASA)-ethyl cellulose systems. The suitability of supercritical carbon dioxide was demonstrated by successfully manufacturing both materials with the minimal required extrusion temperature decreased by 5−20°C.[48,49]

For extrusion granulation, the temperature profile might not be as critical. A simple temperature profile, such as straight temperature profile (ie, all barrels are set at the same temperature) can be considered. However, since the frication factor between the material and barrel wall/liner is influenced by barrel temperature, the conveying efficiency could be negatively impacted by the selected temperature. Usually, the hotter the surface, the lower the friction factor is. Thus, this factor should be evaluated as part of process development to ensure the desired throughput and product quality.

31.3 HOT-MELT EXTRUSION

Interest in HME to prepare ASD of poorly water-soluble drugs began in the 1980s. Since the 1990s, use of HME to manufacture ASDs for improving drug absorption (eg, increase bioavailability and minimize food effect) has increased significantly. This is reflected by the drastic increase in research publications on ASD and the corresponding process technologies since the late 1990s, including chapters in two recently published books.[50,51] It also coincides with the increasing number of BCS class 2 and 4 drugs in new drug pipelines of the pharmaceutical companies.

Amorphous solid dispersions of poorly water-soluble drugs typically consist of an active ingredient, a thermoplastic polymer, a surfactant, and processing aids such as a plasticizer and an antioxidant. Drug loading of ASDs prepared using HME typically is below 40%. In most cases, the carrier polymer is the major component in the formulation. Both a thermoplastic polymer and a surfactant are used to maintain supersaturation and control the precipitation of drug in the aqueous environment. The carrier polymer also prevents the drug from crystallizing during the storage. Processing aids are used to improve the physical and chemical stability of the drug and the polymer during the extrusion process. Depending on their properties and functionalities, individual components in the formulations can be introduced into the extruder at different processing sections. Crystalline drugs and amorphous polymeric carriers are introduced either as a single powder blend or separately into the extruder hopper using loss-in-weight feeders. Liquid additives such as liquid plasticizers and surfactants are added through the liquid injection port. Inside the extruder, the formulation is transformed into a homogenous amorphous material under the intensive mixing of rotating screws at elevated temperatures. After exiting the die, the molten extrudate is air cooled and cut into pellets or cooled with a chill roller and broken into flakes. The pellets and flakes are further milled prior to being incorporated into the final dosage forms.

31.3.1 Formation mechanisms of ASD

During HME, the formulation in the form of a powder blend is transformed to a single-phase homogenous melt under the intensive mixing of the rotating screws and elevated temperature inside the barrel. When the extrudate is cooled quickly to ambient temperature outside the extruder, the solubilized drug is locked in a high-energy state. Even though amorphous API is thermodynamically unstable, a rationally designed ASD can be rendered sufficiently kinetically stability throughout the product shelf life when protected from moisture using an appropriate packaging configuration. The energy needed to form the melt is predominately from the motor and is transmitted by the gear box and

screw elements to the formulation via viscous heat dissipation.

The term *hot melt* does not necessarily imply that the API is always melted during the extrusion process. In fact, depending on the API properties and its interaction with the carrier polymer, HME can be carried out below and above its melting point. When the melting point of an API is below the typical HME process temperatures (< about 200°C) at which the majority of polymers are stable, homogenous dispersion of the API in the molten polymer matrix can be achieved either by API dissolution or due to primarily entropy-driven API-polymer miscibility at high temperature. When the HME processes are conducted above the melting point of the drug substance, homogeneous dispersion can be achieved by mixing the liquid drug melt and the liquid polymer melt.[52]

The intensity of mixing is a function of both the machine parameters (eg, screw geometry) and processing parameters (eg, screw rotation speed, feed rate).[14] The most intense mixing occurs at the sections containing mixing screws. As a result, solubilization of crystalline drugs takes place primarily in those sections. As discussed in the previous section, there are two main types of mixing: distributive mixing and dispersive mixing (see Fig. 31.9). Distributive mixing involves melt division and recombination, while dispersive mixing involves planar and elongational shear. The intensity of mixing also depends on the location of the material relative to the rotating screw and barrel. As shown in Fig. 31.17, material is bounded between screw flights and barrel wall, and separated into small melt pools by screw flights and barrel wall, which provide superior mixing by forming many small mass continuous mixers as compared to the large mass batch mixer. For a bilobal TSE, the high-shear regions, including overflight/tip region, lobal pool, apex, and intermesh, ensure sufficient mixing.

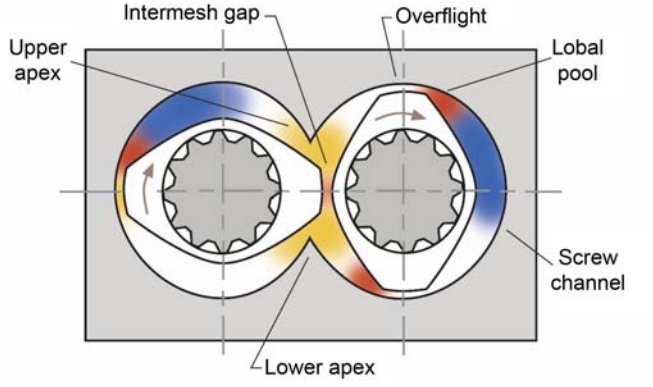


FIGURE 31.17 Cross-section view of bilobal corotating twin screw denoting five shear regions. *Courtesy of Leistriz.*

31.3.2 ASD formulation consideration

As fundamental principles and R&D aspects of amorphous API and ASD systems can be found in chapters "Crystalline and Amorphous Solids," "Solid-State Characterization and Techniques," "Oral Formulations for Preclinical Studies: Principle, Design, and Development Considerations," and "Rational Design for Amorphous Solid Dispersions," this section will focus the discussion on polymers and plasticizers (surfactants) selection consideration.

31.3.2.1 Polymer selection consideration

As polymers typically account for the highest percentage of ASD formulations, their selection is critical. The reason that HME can be conducted below the melting point of a drug is that the melting point of a drug is depressed when there is a strong interaction between the drug and the polymer. Intermolecular interaction between drug and polymer is also important for the physical stability of ASDs. Interaction parameter, a term used to describe the interaction between a drug and a polymer, can be determined experimentally using the melting-point depression method. Good miscibility between drug and polymer manifests itself as the depression of drug melting point in the presence of a polymer. At the depressed melting point (T_M^{mix}), the Gibbs free energy of the crystalline drug must be equal to the Gibbs free energy of the dissolved drug in its saturated solution in polymer melt. Since the Gibbs free energy is a state function, in a sense it only depends on the current state of the system and is independent of how that state has been prepared. We could use a hypothesized liquid drug substance at T_M^{mix} as a reference state. The difference in Gibbs free energy between the reference state and the drug dissolved in polymer melt at T_M^{mix} can be derived from Flory–Huggins equation. The difference in Gibbs free energy between the reference state and the crystalline drug at T_M^{mix}can be derived using thermodynamic parameters of pure drug substance. Therefore, the following equation is derived to describe the relationship between the depressed melting point of a drug substance and the interaction parameter between a drug and a polymer.[53]

$$\frac{1}{T_M^{mix}} - \frac{1}{T_M^{Pure}} = -\frac{R}{\Delta H_{fusion}} \left[Ln\varnothing_{drug} + \left(1 - \frac{1}{m}\right)\varnothing_{polymer} + \chi\phi_{polymer}^2 \right] \tag{31.9}$$

T_M^{mix} is the depressed melting point of drug in the presence of polymer. T_M^{Pure} is the melting point of pure

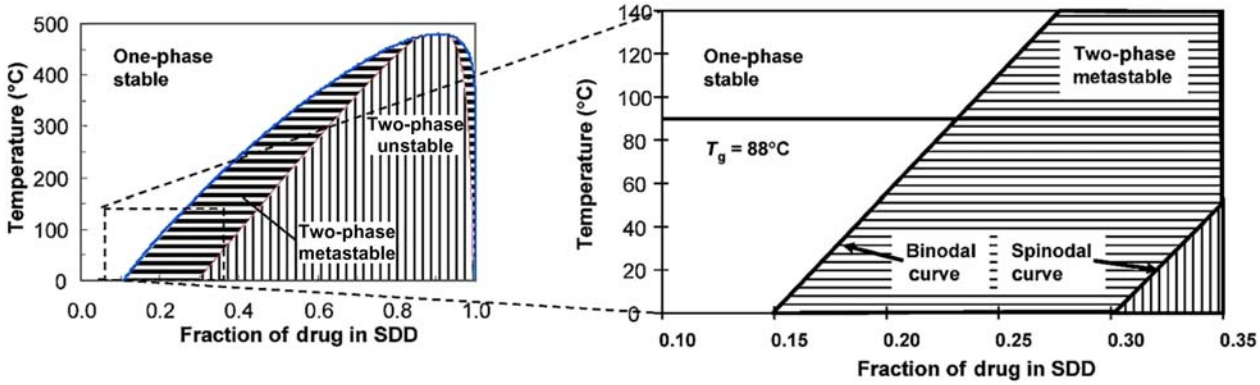


FIGURE 31.18 Phase diagram of drug substance-polymer solid dispersions.[56]

drug substance. ΔH_{fusion} is the melting enthalpy of pure drug substance, ϕ is the volume fraction, m is the polymer-to-drug molar volume ratio, and χ is the drug-polymer interaction parameter.

Once the solubility parameter becomes available, the drug-polymer phase diagram can be readily constructed using the Flory–Huggins equation. An example phase diagram is presented in Fig. 31.18.[56] The phase boundary where $\Delta G_{mix} = 0$ is defined as the binodal curve. A drug solution in polymer is thermodynamically stable outside the bimodal curve. As drug concentration is further increased, the spinodal curve is reached. The spinodal curve is defined as the line on which the second derivative of Gibbs free energy of mixing is zero. Composition between the spinodal and bimodal curves is metastable because large fluctuations are required for the system to phase separate into drug-rich and polymer-rich phases. Within the region surrounded by the spinodal curve, the composition is the least stable, and even a small fluctuation leads to phase separation (spinodal decomposition). The polymer must be thermoplastic so that it could be processed using HME. During the early stage of formulation development, selection of the polymer is based on its ability to maintain supersaturation and control the precipitation of drug in aqueous medium and to maintain physical stability of ASDs during the storage. It is important to realize that the melt extrusion process conditions are dictated to a great extent by the properties of polymers.

In most cases, there are specific interactions such as hydrophobic and dipole interactions and entropy-driven miscibility at high temperature between drug and polymer. Therefore, it is challenging to theoretically calculate the interaction parameters between drug and polymer in ASD systems. Instead, a common practice is to measure the melting point depression and then fit the data to derive the interaction parameters using Eq. (31.9).[54,55] As drug concentration is further increased, the spinodal curve is reached. Spinodal curve is defined as the line where the second derivative of Gibbs free energy of mixing is zero. Composition between the spinodal and bimodal curves is metastable because large fluctuations are required for the system to phase separate into a drug-rich and a polymer-rich phases. Within the region surrounded by the spinodal curve, the composition is the least stable, and even small fluctuation leads to phase separation (spinodal decomposition). A sample drug-polymer phase diagram of an ASD is presented as Fig. 31.18.

The degradation temperature and viscoelastic properties of polymers are the two most critical attributes that influence the extrusion conditions. Degradation of the polymer can be characterized using thermogravimetric analysis, and it can be divided into main-chain reaction and side-chain reaction. The main-chain reaction comprises the chain scission and cross-linking, while the side-chain reaction comprises the side-chain elimination and cyclization. The viscoelastic properties of the polymer must be investigated in terms of the torque limitations of extrusion processing.[57] The viscoelastic properties of a polymer can be described in terms of its storage modulus (G′, how elastic or solid-like the polymer is) and its loss modulus (G″, how viscous or liquid-like the polymer is). The loss tangent is the ratio of the loss modulus to the storage modulus (G″/G′), and a value of 1 represents the transition point from solid-like to liquid-like properties. The sum of the storage modulus and loss modulus is the complex modulus, which represents the overall resistance to deformation.

Pharmaceutical polymers commonly used in melt-extruded ASDs could be categorized into three different groups: polyvinyl-based polymers, cellulose-based polymers, and polymethacrylate-based polymers. These polymers are summarized in Table 31.6. Most of

TABLE 31.6 Common Polymers for Amorphous Solid Dispersions Prepared Using Hot-Melt Extrusion

Polymer	T_g (°C)	Grades	Notes
Hypromellose	170–180	Methocel E5	Nonthermoplastic API must plasticize Excellent nucleation inhibition Difficult to mill
Hypromellose acetate succinate	120–135	AQOAT-L AQOAT-M AQOAT-H AquaSolve-L AquaSolve-M AquaSolve-H	Easily extruded without plasticizer Process temperatures >140°C Ionic polymer soluble above pH 5.5 depending on grade Excellent concentration enhancing polymer Stable to 190°C depending on processing conditions
Modified hypromellose	117–128	Affinisol 15cp, 100cp, and 4 M	Designed for melt extrusion Broad processing window API plasticization not required Excellent nucleation inhibition
Vinylpyrrolidone	168	Povidone K30	API must plasticize Potential for H-bonding Hygroscopic Residual peroxides Easily milled
Vinylpyrrolidone-vinylacetate copolymer	106	Kollidon VA 64	Easily processed by melt extrusion No API plasticization required Less hygroscopic than povidone Processed around 130–180°C
Polyethylene glycol, vinyl acetate, vinyl caprolactam graft copolymer	70	Soluplus	Designed for melt-extruded dispersions Easily process by melt extrusion Low T_g can limit stability Not of compendial status Stable up to 180°C
Polymethacrylates	130	Eudragit L100-55 Eudragit L100	Not easily extruded without plasticizer Degradation onset is 155°C Ionic polymer soluble above pH 5.5

these materials were originally developed as binder or coating materials in conventional manufacturing processes. In recent years, novel excipients have been developed specifically for the melt-extrusion process. They typically have a wider temperature window for extrusion and are more chemically stable during the extrusion. For instance, BASF became the first company to engineer a polymer for the melt-extrusion process when Soluplus was introduced in 2009.[58] Soluplus is an amphiphilic polyvinyl caprolactam-polyvinyl acetate-polyethylene glycol graft copolymer. Dow Chemical recently introduced Affinisol, a new grade of hypromellose specifically designed for the melt-extrusion process. While conventional hypromellose E5 has been difficult to extrude due to high T_g and low thermoplasticity, the new line of Affinisol was developed with a lower melt viscosity, reduced discoloration at elevated temperatures, and wider processing window.[59] SE Tylose has developed a new grade of hypromellose acetate succinate (HPMCAS) with a lower glass transition temperature of 100°C. By comparison, the current commercial grade of hypromellose acetate succinate has a glass transition temperature of 120–135°C and must be processed at a temperature above 160°C.

31.3.2.2 *Plasticizer and surfactant selection considerations*

The addition of a plasticizer to facilitate the melt-extrusion process is well established in the plastics and pharmaceutical industries. Plasticizers are able to lower the glass transition temperature and the melt viscosity of polymer. For example, 5–10% triethyl citrate was required in the formulation to minimize degradation of troglitazone when Resulin was developed. Surfactants such as sorbitan monolaurate and vitamin E TPGS are present in ASDs of several commercial products. These surfactants are present to modify the precipitation and crystallization behavior of ASDs in aqueous environment. These surfactants have demonstrated plasticizing effects.

The major concern with using the traditional plasticizers is that the ASDs' physical stability might be compromised during the storage because of the lower glass transition temperature. This problem may be addressed by using transient plasticizers.[48,49] A transient plasticizer is intimately mixed with the polymer melt inside the twin-screw extruder and subsequently removed completely from the extruded product upon exiting the die. Supercritical CO_2, a blowing agent commonly used in plastics industry, is an ideal transient plasticizer for pharmaceutical melt extrusion. A schematic of a supercritical injection system coupled with a twin-screw extruder is shown in Fig. 31.19. The chiller and pump assist in maintaining CO_2 in the supercritical fluid. Studies conducted at Merck have demonstrated that the supercritical CO_2 can lower the processing temperature by as much as 20°C.

31.3.3 HME process consideration

As mentioned earlier, extrusion is an integrated unit operation. HME typically consists of several functional zones: (1) feeding, (2) solid conveying, (3) mixing/melting, (4) devolatilization, (5) melt conveying, and (6) melt discharge. Individual process parameters, including feed rate, screw speed, barrel temperature profile, and vacuum, are adjusted independently to meet product quality requirements. Refer to Section 31.2.2 for the general guidance on process design. As mentioned in prior sections, screw configuration should also be properly designed to achieve product requirements. Common screw elements that can be used for extrusion have been discussed in Section 31.2.1.2. A screw configuration, which is achieved by strategically placing different screw elements at different functional zone, ensures the desired mixing and pressure profile. The pressure gradient along an example screw in a starve-fed TSE is shown in Fig. 31.20. As discussed in Section 31.2.1.2, different screws typically serve different functions. An example screw design is presented in Fig. 31.21. In this case, conveying screws are used to convey solid and melt. Mixing or solubilization of drug in polymer melt is enabled using kneading elements, and mixing screws are used to create melt seal before devolatilization (Fig. 31.22).

One of the most challenging aspects in developing an HME process for manufacturing ASD is to achieve a balance between obtaining uniform ASD formulation by providing sufficient mixing while minimizing degradation of drug and/or polymer. In process design, proper selection of screw design and process

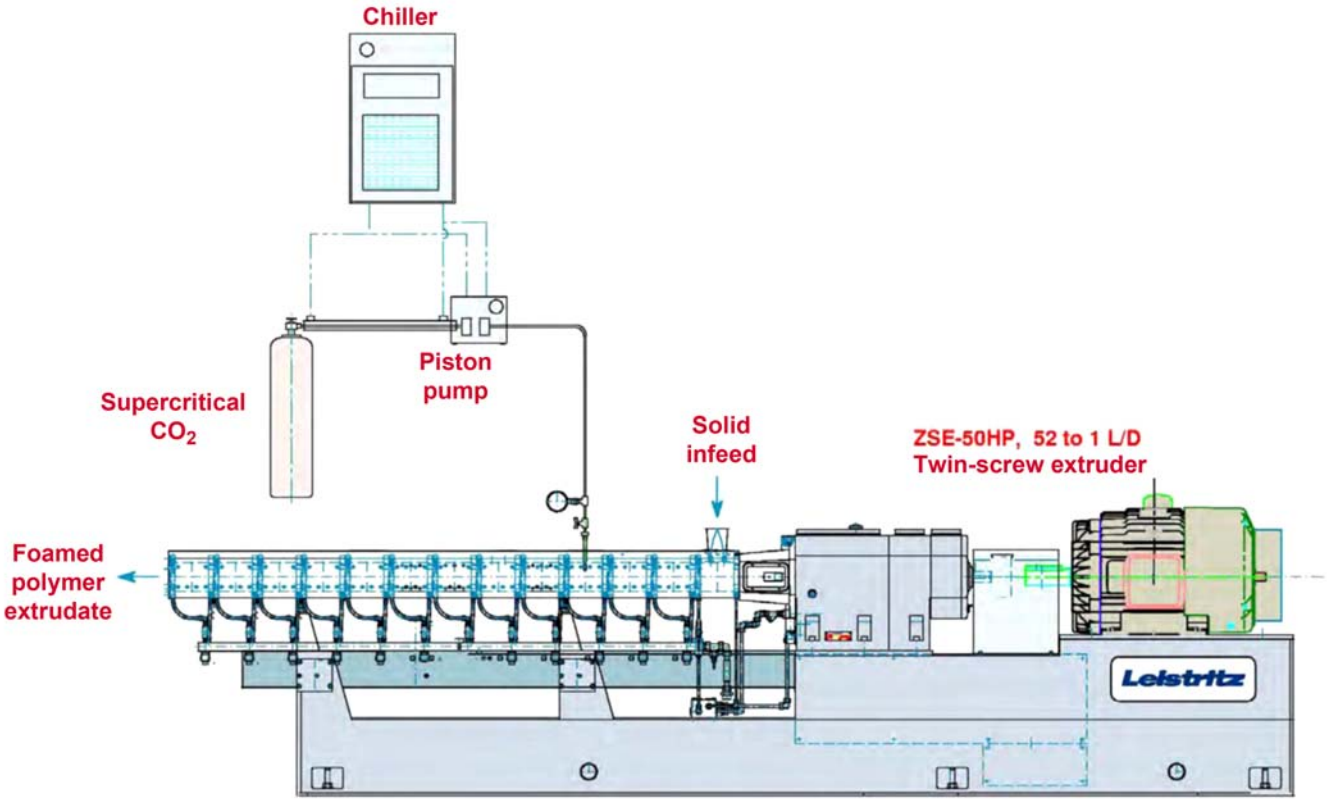


FIGURE 31.19 Schematic diagram of injection of supercritical CO_2 during extrusion process. *Courtesy of Leistriz.*

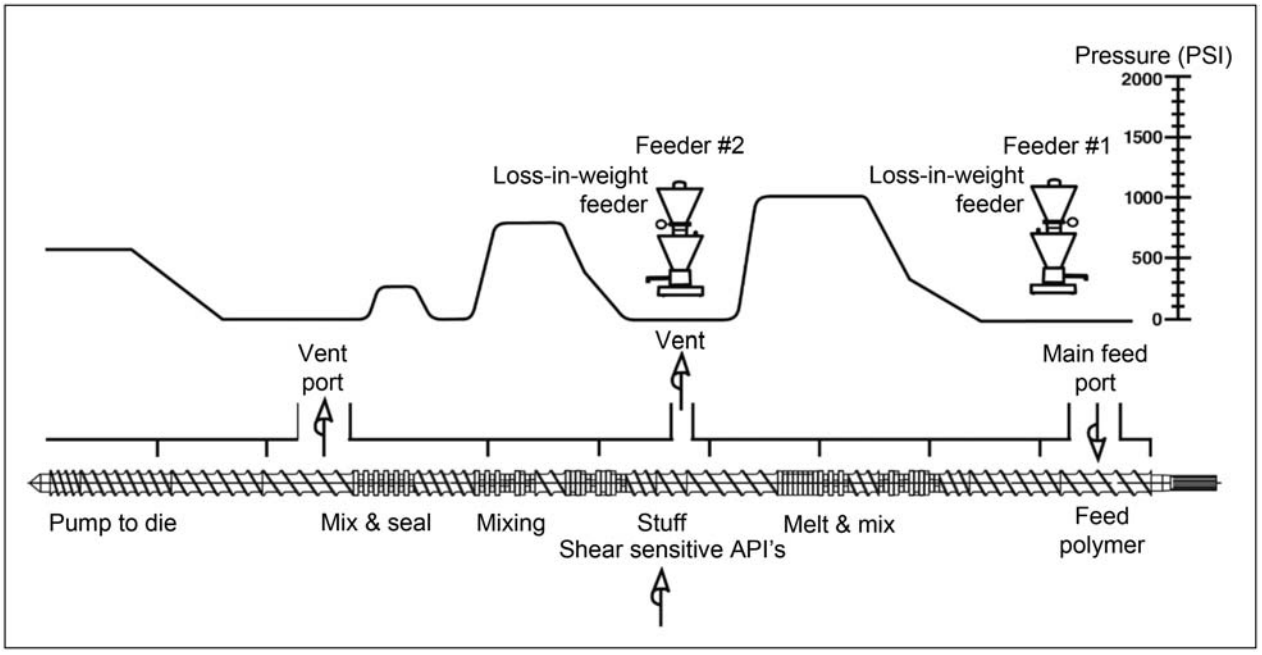


FIGURE 31.20 Pressure gradient in a starve-fed TSE process. *Courtesy of Leistriz.*

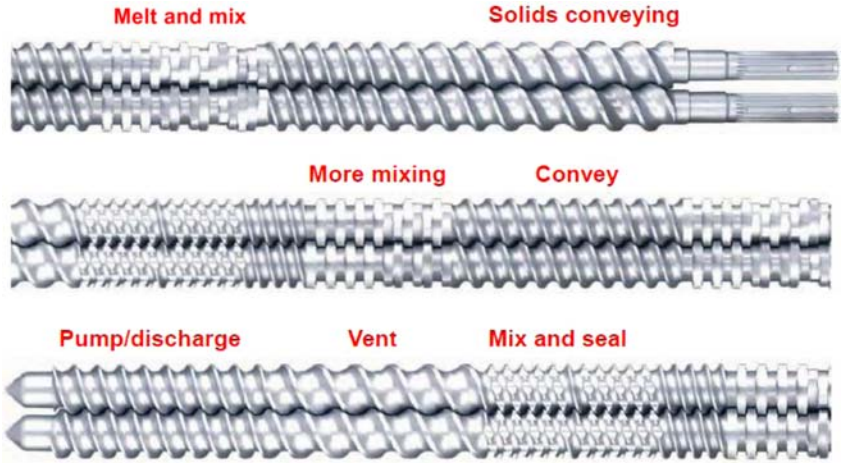


FIGURE 31.21 Different types of screw designs for different types of mixing operations. *Courtesy of Leistriz.*

FIGURE 31.22 Different screw designs used to evaluate indomethicin-eudragit E ASD.[61]

parameters including barrel temperature profile, feed rate, and screw speed is critical, as these extrusion variables control the key systems parameters (eg, shear rate, degree of fill, residence time, residence time distribution, and melt temperature) and thus affect product quality, as shown in Fig. 31.15. The impact of each process parameter is discussed in Sections 31.2.1.2 and 31.2.2. For compounds prone to thermal and/or shear degradation, the degradation occurs primarily at the mixing/melting zones, where the formulation is exposed to the highest shear and thermal stress. In these zones, the local melt

temperature is predominately controlled by the viscous dissipation as the result of the shearing effect by the rotating screw.

Screw configuration is an important parameter for HME. Altering the screw configuration allows for the modification of the shear stress and residence time. Different screw elements can be optimized to suit particular applications. A right screw design should be able to provide sufficient mixing capacity with the proper shear stress. Generally, one kneading zone positioned at the two-third or at the end of the barrel is adequate to create sufficient mixing at the molecular level for generating ASDs. Caution should be taken when considering modifying the kneading elements, as varying kneading elements could significantly affect the product quality due to a change in residence time and shear stress. Nakamichi et al. concluded that at least one mixing zone-containing kneading element was needed to obtain homogeneous extrudates of nifedipine (NP) and hydroxypropyl methylcellulose phthalate (HPMCP).[5] Gosh et al. demonstrated that the position of the kneading block is critical for processing of thermally labile compounds. Higher drug degradation was observed when the kneading block was close to the feed zone. When kneading block is moved closer to the end of the barrel, API stability improved due to delayed melting and shorter melt residence time.[60] Liu et al. evaluated the effects of four screw configurations on miscibility behavior of indomethacin (IND) in Eudragit EPO. They concluded that IND could not be dissolved into the Eudragit EPO matrix efficiently using screw configuration A (ie, without mixing/kneading zone). They also indicated that the first kneading zone accelerated the dissolution of IND into Eudragit EPO, while the subsequent second kneading zone was superfluous.[61]

Material properties (T_g, T_m, miscibility/solubility regime) should be taken into account when selecting barrel temperature as it would affect the melting process. The heat conducted from the barrel usually assists the melting process at the beginning of the process. However, it might not play a significant role in melting materials at steady-state of the extrusion process. In fact, when additional dispersive and/or distributive mixing is present, the melt temperature is often much higher than the barrel temperature. Under such circumstance, the heated barrels actually act as a chiller to remove the excessive heat generated by shear. Even if conductive heating is not critical to material melting, it still affects the melt viscosity. Increasing the barrel temperature decreases the melt viscosity and thus decreases shear stress (as shown in Eq. 31.4), which might reduce thermal degradation. When the melt temperature is higher than the barrel temperature, increasing the barrel temperature lead to high melt temperature and thermal stress due to decreased barrel cooling capacity. In this situation, degradation formation is ultimately determined by the combined effects from both variables. If the increase in thermal stress is greater than the reduction of shear stress, the degradation could be accelerated with increasing barrel temperature. Jijun et al. have shown that maintaining barrel temperatures above 130°C is required for producing 15% nimodipine amorphous dispersion in copovidone.[62] When the barrel temperature rises above 160°C, degradation of nimodipine was observed with a color change of the extrudate from gold to brown. Thus, it is important to note that the melt temperature is consequently the most critical determining factor in the product quality, and it might be very different from the barrel temperature. In some cases, the barrel temperature might have little influence on product quality. Lyons et al. demonstrated that barrel temperature does not affect either the inclusion of API in the polymer matrix or in-vitro drug release from carvedilol and polyethylene glycol-*co*-polycaprolactone (PEO−PCL) matrices within a temperature range of 95−120°C.[63]

For a given screw design, increasing screw speed increases shear stress and specific energy input to the materials while decreasing residence time. A minimum screw speed is required to ensure sufficient specific energy and residence time required for solubilizing the drug in the polymer melt. The mean residence time for HME process is typically less than 60 seconds. However, excess shear/specific energy resulted from higher screw speed or extended residence time resulting from lower screw speed could negatively affect chemical stability. Thus, when defining a screw speed range, all of these factors should be evaluated carefully. Crowley et al. found that when processing PEO with chlorpheniramine, a low screw speed with a longer residence time induced polymer degradation due to prolonged exposure of PEO to the extrusion temperature. Increasing screw speed led to reduced melt viscosity and identification of the optimal processing condition. However, polymer degradation occurred with further increase in the screw speed due to the heat generated by the mechanical energy of the screws.[64] Gosh et al. reported similar results on processing a thermally labile API. A low screw speed induced thermal degradation due to a longer residence time, while a high screw speed caused degradation due to increased mechanical energy.[60] For a starve-fed extruder, feed rate also impacts residence time and shear stress input to the materials by affecting degree of fill. To maintain a constant melt flow, a balance needs to be identified between the screw speed and the feeding rate. Under constant screw speed, higher feed rate results in shorter residence time. Dependent

on the selected screw configuration, specific energy might increase or decrease with increasing feeder rate. Therefore, HME process design should focus on the understanding of the key system parameters (ie, the combination effects) rather than individual process parameters.

The stability of an ASD formulation could also be impacted by the temperature rise resulted from the pressure generation in the extruder front-end. The more restrictive the front-end, the higher the melt pressure and temperature, which may adversely affect the product quality. The temperature rise can be estimated by Eq. (31.10).

$$\Delta T\ (^{\circ}\mathrm{C}) = \frac{\Delta P\ (bar)}{2} \tag{31.10}$$

ΔT is the change in temperature as the result of melt pressure buildup. ΔP is the melt pressure (1 bar = 14.503 psi).

It should be pointed out that use of an HME process for manufacturing ASD is not without limitations despite its many advantages. For ASD-containing drug substances with a high melting point (eg, >220°C), limited miscibility with the carrier polymer matrix, or both, formation of an ASD can be more challenging, and low drug loading is often observed when compared with other processing technologies, such as spray-drying. Therefore, rational design and development of ASD formulation and HME process requires integrated consideration of the basic properties of the individual API, polymers, surfactants, processes and their interplays as well as the required dose. There is no one-size-fits-all formulation or processing technology for developing and manufacturing ASD products.

31.4 CONTINUOUS GRANULATION USING A TWIN-SCREW EXTRUDER

Granulation is a process in which small particles are combined to form larger permanent particles. The granulation process is widely used in the pharmaceutical industry to improve the flow, density, uniformity, and compressibility of the material for downstream processing. Granulation methods can be categorized into wet granulation, dry granulation, and melt granulation. Wet granulation utilizes liquid to bind the primary particles. High- and low-shear mixers and fluidized bed granulators are typically used in a traditional batch process for the wet-granulation process. Dry granulation does not require the addition of any liquid and solely depends on mechanical stress to bind the particles together. Roller compactors and sluggers are typically used. Both methods are operating in a continuous mode. In the melt-granulation process, granules are formed through the softening or melting of binders. High-shear or fluid-bed granulators are traditionally used for melt granulation in a batch mode.

With increasing interest in continuous manufacturing, the feasibility of continuous wet and melt granulation is revisited by researchers from academia and industry. The advantages of continuous granulation over conventional batch granulation include: (1) reduced cost, resources, and time related to scale-up for product launch; (2) less material handling; (3) more streamlined and efficient process; and (4) more robust process and consistent product quality.

Extrusion, fluid-bed agglomeration, instant agglomeration, and spray drying have been used as the techniques for continuous granulation.[65] Among these techniques, extrusion has been the most popular for pharmaceutical applications. Gamlen and Eardey (1986) first reported using a twin-screw extruder for a granulation process in 1986.[66] This technique attracted more attention after a patent for twin-screw granulation for application in a single-pass, continuous pharmaceutical granulation process in 2002.[122]

Continuous granulation via a twin-screw extruder can be used to replace the traditional granulator in a bin-to-bin manufacturing setup, or it can be integrated into a continuous manufacturing line as demonstrated in GEA's ConsiGma system. For a conventional tablet manufacturing process using a wet granulation step, it contains the following unit operations: blending of the raw materials, granulation, drying, milling, blending of extragranulate, and tableting. Using a twin-screw extruder can potentially eliminate the blending of intergranulate by feeding the ingredients using loss-weight feeders directly to the extruder. With proper screw design and process setup, unimodal particle size distribution (PSD) of granules might be achieved without a milling process for some formulations, which can significantly streamline the process. The more efficient mixing in the twin-screw extruder also requires less binder solution for wet granulation, allowing the drying time to be reduced. For melt granulation using a twin-screw extruder, high efficiency of mixing with more consistent temperature control and significantly lower residence time allows a wider range of process temperature than a normal batch melt granulation process.

As discussed in Section 31.2, a twin-screw extruder consists of two intermeshed screws enclosed in a barrel that rotate in the same direction in most of the applications. The extruder used in HME described in Section 31.3 is often utilized for continuous wet or melt granulation with modified equipment and process parameters, such as screw configuration and temperature profile. A schematic drawing of the standard

layout is shown in Fig. 31.23. Multiple loss-in-weight feeders can be used to feed different ingredients at the same time as needed. Liquid (solvent, water, binder solution, or molten binder) can be fed through a liquid injection system into the extruder. Different liquid feeding positions can be explored depending on different formulations, and the barrel temperature can be controlled. In wet granulation applications, the barrel is normally maintained at room temperature. In melt granulation applications, the barrel can be heated to maintain the temperature required to keep the binder at molten state. The modular design of the screw elements provides the flexibility to meet the product design and processing requirements.

31.4.1 Continuous wet granulation

31.4.1.1 Granulation mechanism inside extruder

For the continuous wet granulation process using a twin-screw extruder, as shown in Fig. 31.23, a powder blend is fed into the beginning of the extruder. Binders can be fed individually using a loss-in-weight feeder, in a powder blend with other ingredients, or as a predissolved solution or foam. There are multiple positions for liquid injection along the extruder. Process parameters, such as feed rate, screw speed, liquid-to-solid ratio (L/S ratio), and screw configurations, together with the formulation design determine the granulation properties. A fundamental understanding of the granulation mechanism inside the extruder helps determine the combination of different process parameters and formulation variables, therefore, enabling the efficient design of a continuous granulation process to achieve the desired product properties.

Different from the granulation process in a high-shear or fluid-bed granulator, wetting, nucleation, agglomeration, and breakage may happen in sequence along the extruder in a much shorter time within a twin-screw extruder.[67] Nucleation happens when the binder solution reaches the powder bed during a granulation process. The method by which the binder solution is applied (by spraying vs dropwise addition) can change the droplet size of the binder solution and, therefore, determines the mechanism of granulation. When the binder solution is fed as liquid drops into the extruder through a liquid nozzle, the droplet size is usually much larger than the powder particle. The immersion mechanism dominates, transforming the powder particles into big, wet, loose lumps.[68] In most cases, this nucleation phase takes place in the conveying screw element region. When the agglomerates reach the kneading elements, they are sheared, compressed, deformed, and fractured. The binder solution is redistributed in this region (eg, being squeezed to the particle surface).[69] Consolidation of granules occurs, and the granules become denser and stronger. Some large granules might even be fractured into smaller granules.

Granulation liquid mixing and distribution during continuous granulation can be visualized by NIR chemical imaging.[70] The distribution of the granulation liquid in the granule was used as an indicator to the granulation formation. The distribution of granulation liquid was more homogeneous at higher L/S ratio. However, when moisture homogeneity is dramatically increased, the PSD of granule remained bimodal. This finding suggested that the typically observed wide and bimodal PSD from a TSE granulation is inherent to the process rather than caused by the insufficient mixing of powder and liquid. Though milling can be used to alter and control the particle size and its distribution for tableting, more fines could be generated during milling, increasing the segregation tendency during subsequent material handling and transfer for downstream processing. Therefore, efforts have been taken to use the modular design of the extruder to produce granules with the desired properties. The kneading block, conveying, mixing, and tooth-mixing elements can be added or combined to the design as modules. There are studies that have shown unimodal

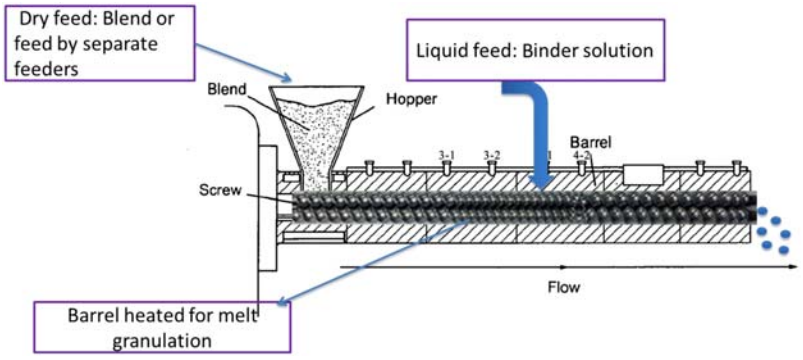


FIGURE 31.23 Schematic drawing of a twin-screw extruder used in continuous granulation.

distribution of granule particle size under certain conditions.[68,71] To help further understand the granulation process using TSE, the effect of screw design and process parameters is discussed in the following sections.

31.4.1.2 *Effect of screw design on granulation*

Three basic types of screw elements used for granulation include conveying screw, kneading, and DMEs. Detailed description of these screw elements can be found in Section 31.2.1.2. The modular design of screw elements provides flexibility in the process design. However, it can make the process development challenging due to the extensive number of available combinations and permutations. Thus, understanding the basic working principles of different types of screw elements and their impact on granulation can help with rational selection of the screw elements to aid the development of a TSE granulation process.

31.4.1.2.1 Conveying

Granule formation in conveying element is mainly due to particle collision. The granule formed is weak and can be fractured into fines easily through the friction force against the barrel wall or towards the flight of the screw. As a result, a significant amount fines are often observed. Thompson and Sun studied the evolution of agglomerates inside the extruder with only conveying elements.[4] The fraction of fines decreased along the length of the screw and mostly over the latter half of the screw. More fines were generated at a low-fill level compared with a high-fill level due to more frequent collisions between particles to facilitate granule growth and smaller portion of particles that are subject to the high-shear force at the barrel walls. Granules are mostly oblong shaped.

Conveying elements are normally used in combination with a kneading element to form stronger granules. The impact of a conveying element before and after a kneading element on granule properties can be very different.[72] Liu et al. (2015) studied the impact of the conveying screw element before and after the kneading zone on granule properties using a ZSE-HP 27 mm extruder. Three values of flight pitch (P/D = 0.67, 1 or 1.33) before and after a kneading block and two different degrees of fill level were evaluated using a microcrystalline cellulose/ lactose system and a system with the addition of ibuprofen. Only the conveying element after the kneading zone had a significant impact on granule properties.[72,73] The screw pullout picture shown in Fig. 31.24 reveals the changes in the granule characteristics across the different zones when the materials were extruded with different screw configurations at a feed rate at 15 kg/hour. The circled region is where the particles exit from the kneading block. Three different screw configurations are

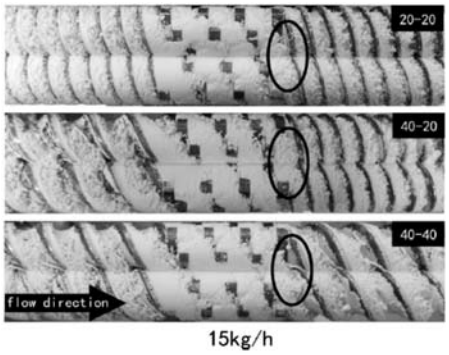


FIGURE 31.24 Screw pullouts for three screw configurations showing the granulation progress along the screws in the region of the kneading block (feed rate = 15 kg/h).[72]

demonstrated. The one on the top uses conveying elements with a flight pitch of 20 mm before and after the kneading block. The one in the middle uses a conveying element with a 40-mm flight pitch before the kneading block and conveying elements with a 20 mm flight pitch after the kneading block. The one at the bottom uses conveying elements with a flight pitch of 40 mm before and after the kneading block.

In general, collision-dominated agglomeration occured in the upstream conveying zone. The binder distribution varied considerably among the different particle sizes due to the complex velocity profiles in the conveying zone. Granule size at this conveying zone increased with the increase of the flight pitch. Velocity of particles conveyed forward increased with the increase of flight pitch, which led to more particle agglomeration. Increasing flow rate reduced the size and shape variation of the granule. This was due to a relatively reduced frequency of collisions in a wider channel. Porosity and strength of the granules were not impacted by flight pitch and flow rate. A downstream conveying element contributed to the fragmentation of compacted particles exiting the kneading block. After the kneading block, the compacted particles were extruded towards the flight of the downstream conveying element. The impact of downstream conveying element on the granule particle size was smaller when a wider flight pitch conveying element was used due to the larger angle between the obstructing planar walls. Larger, more elongated, and fragile granules were observed in the larger pitch conveying element.

31.4.1.2.2 Kneading element

Granule size and strength are mainly determined by the kneading element. Granule agglomeration and breakage occur simultaneously in this region. A kneading block comprises a number of discs, each offset by an angle from the next. The offset angle is the main variable varying the drag/pressure flow capacity of

a kneading block without changing the available volume. The disk thickness changes the volume of material flow in the kneading block. At a lower offset angle, drag flow dominates over pressure-driven flow. With increasing offset angle, the pressure-driven flow increases until only pressure-driven flow occurs at a 90° offset angle. At high screw speed (low-fill level in the conveying zone), the offset angle does not seem to affect the granule particle size. The PSD of the granules only changes with more large agglomerate at low screw speed when a 90° offset angle kneading block is used.[4] This phenomenon is very different from the HME process, in which the offset angle of a kneading block can significantly influence the extensional flow as a fluid is squeezed through the gaps of the offset disks. For cohesive granule flow, granule development appears to occur in the intermeshing region, and the offset angle of the kneading disks plays only a minor role in the compressive stress in that zone. The granules from the kneading block presents plate or ribbon-like granules due to the calendaring action when the granule is compressed through the adjacent rotating kneading disks in the intermeshing region. For example, a long reverse-flighted kneading block has been reported to cause blocking during granulation.[74]

31.4.1.2.3 DMEs

Broad bimodal size distribution, with the presence of lumps and ungranulated fines, are observed for the granulation using a twin-screw extruder in many cases. Drying uniformity of these granules is often difficult to control. To address this issue, a DME has been explored to alter the granule size distribution.[71] A picture of a DME is shown in Fig. 31.25. The angularly cut blades are perpendicular to an annulus portion of each DME. The annulus portion is thicker than the blade, as shown in the side view. The portion extending out the blade is known as the spacer. A DME can be mounted in two ways depending on the position of the spacer (see Fig. 31.26). If the spacer is placed facing upstream, it is called the *forward configuration*. If the spacer is placed facing downstream, it is called the *reverse configuration*.

A DME allows the mixing and collision of granule particles through cutting and recombination. Like the kneading block section, the DME section is also a mainly pressure-driven flow section.[4] The forwarding comb element has similar performance as the kneading block, which produced coarse granules, but a considerable amount of fines is still observed, which is likely due to the angular cuts by the mixing element. However, the granule particle shape generated is different from a kneading block. Round-to-oblong shaped granules are formed using the comb mixing element instead of elongated plate-like granules generated using

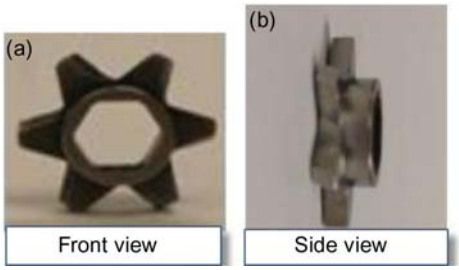


FIGURE 31.25 Pictures of a DME.[71]

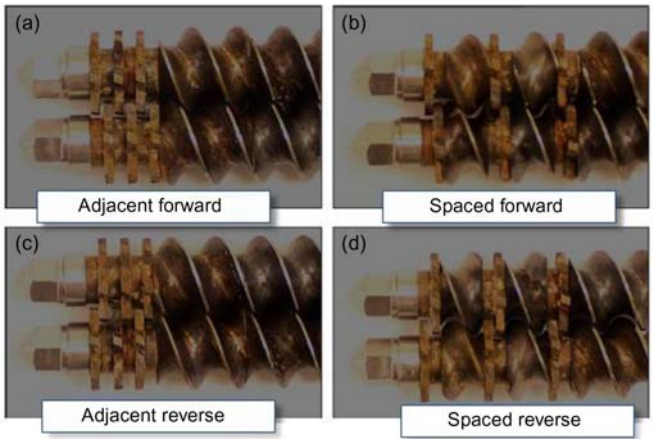


FIGURE 31.26 DME configurations.[71]

the kneading element. The particles are packed intensively in the annular region between adjacent rings. When the packaged granules pass the intermeshing region, they are chopped by the segmented rings of the adjacent comb mixing element. Granule formation can also occur in the gaps of the rings. A reverse-comb element produces complete agglomeration, which often leads to blockage, and is not recommended.

Sayin et al. studied four different configurations (see Fig. 31.26) of DMEs using a 16-mm Eurolab twin-screw extruder. Breakage and layering are the main mechanisms of granule formation when a DME is used. Large agglomerates from the convey section were broken by the rotating blade of a DME followed by layering of ungranulated fines on the newly exposed wet surfaces. Granule particle size was more significantly impacted by the orientation (spacer facing upstream or downstream) of the DME than the placement (regardless the presence of the conveying element between DME). The reverse configuration resulted in narrow PSD and more uniform liquid distribution compared to the forward configuration. In the forward DME configuration, the bimodal distribution of granule particle size often persisted until the liquid-to-solid ratio was very high, where agglomeration of the fines could take place more easily. In the

reverse DME configuration, unimodal distribution of granule particle size could be achieved at much lower liquid-to-solid ratio. The granule property difference observed was due to the difference in the spacer position in the forward DME and reverse DME. In the forward DME configuration, the spacer was between the blade of DME and helix of the conveying element. Materials were staggered between this space and carried over along the internal barrel wall without being chopped by the blade; therefore, the material flow was homogeneous and continuous. In the reverse DME configuration, the material flowed to the blade directly and was chopped.

The impact of the number of DMEs on granulation was studied by Vercruysse et al. (2015)[75] When a DME was used, the amount of fines was significantly reduced compared to the conveying-only configuration. An increase in the number of narrow DMEs (thinner blade, Fig. 31.27) did not further reduce the amount of fines; however, the additional DMEs could reduce the amount of larger granules. The addition of a third narrow DME did not have any impact on the particle size. Due to nonforwarding characteristic of the DMEs, the use of more than one wide DME (wider blade, see Fig. 31.27) is not recommended because it resulted in high and unstable torque.

FIGURE 31.27 Pictures of wide DME and narrow DME.[75]

31.4.1.2.4 Screw mixing element

The screw mixing element, as shown in Fig. 31.28, can cause more backflow, resulting in longer residence time compared with traditional conveying elements.[75] Screw mixing elements do not generate high-shear regions as the kneading element due to their higher transport capability. The granule PSD obtained using a screw mixing element is very similar to the one from conveying elements, with a high percentage of fines and a large portion of extremely large agglomerates. Comparing with screw configuration consisting only of conveying elements, a screw mixing element can slightly reduce the amount of fines. More screw mixing elements are usually required to decrease or avoid extremely large agglomerates.

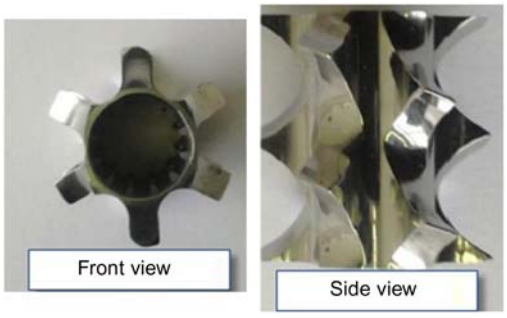


FIGURE 31.28 Pictures of a screw mixing element.[75]

31.4.1.2.5 Screw configuration

The screw configurations for the wet granulation process should always start with conveying elements, followed by kneading or DMEs. In the conveying zone, the fill level is low, and particle densification and mixing are typically very limited. Granules are formed through powder layering on the liquid droplet. There is no redistribution of the liquid in the granules; therefore, the uniformity of the liquid distribution is very poor. Multimodal PSD with a high amount of fines and a small portion of agglomerates are often obtained. Only at high L/S ratios can more agglomeration take place, resulting in fewer fines and more agglomerates. When the large agglomerates formed in the conveying zone reach the kneading zone where the fill level is significantly increased, shearing, compressing, and deformation occur, leading to densification of the material, significant reduction in the amount of fines, and an increase in the agglomerates. At a high L/S ratio, unimodal PSD can be achieved; however, the granule particle size is normally too big to be used directly for tableting. Similar results were observed when using a DME.[71] When two kneading zones were used, the second kneading zone broke down the big agglomerates formed in the first kneading zone by shearing the overwetted mass between the tips of the kneading element and the barrel wall. At the same time, the fines were exposed to the newly formed wet surface of the broken agglomerate. Therefore, when two kneading zones

were used, the portion of fines and the extreme big granules were reduced, and the median particle size was increased. When the number of kneading segments is increased, more large agglomerates were formed. The liquid distribution in the kneading zone was much better compared to the conveying zone due to the dispersive mixing. Improved uniformity was achieved with a longer kneading zone. In the case of two kneading zones, a homogenous distribution can be achieved.[4] Alternatively, DMEs can be used to replace the kneading elements or can be combined with additional kneading elements to control the granule particle properties. The screw elements at the exit are important as well.[75] When two narrow kneading elements (length = D/6 for each element) at a 30° angle were used, no change was observed in granule PSD compared with the use of the convey elements. When a wide kneading element (length = D/4 for each element) was used, further agglomeration takes place. Particles were forced into the gap of the kneading element and the barrel wall. Increasing the number of wide kneading elements used resulted in more agglomeration. However, when the offset angle was 60° or 90°, no difference in granule PSD was observed compared to the conveying element, due to the larger axial opening between the kneading disks. The use of cutter elements (see Fig. 31.29) led to more oversized agglomerates. In addition, it increased the pressure at the end of the screw chamber, giving rise to high and variable torque profiles. When a single narrow DME was used, a reduction of particles larger than 1000 micron was observed. An additional DME further reduced the large agglomerate, but it introduced more fines. It was found that the combination of one kneading block and a DME at the end seemed to optimize the PSD. More DMEs at the end reduced the median particle size.

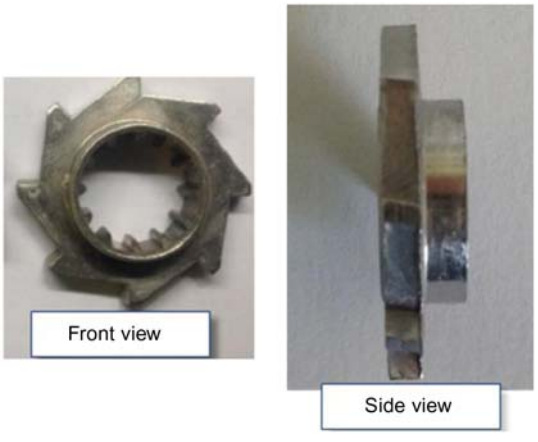


FIGURE 31.29 Pictures of cutter.[75]

31.4.1.3 *Effect of extrusion process parameters on granulation*

In addition to screw elements, feed rates (solid/liquid ratio), screw speed, and liquid injection position are other parameters that can be adjusted to change the granule properties. The effects of these process parameters on granulation also depend on formulation compositions. Some materials or formulations are more robust, and the resulting granule properties do not vary significantly with the change in process parameters. In other cases, the granule properties can change dramatically with a change in the process parameters. For example, lactose is not as sensitive to shear and over-wetting during granulation when compared to dicalcium phosphate.[76,77] In this section, the influence of process parameters on the granulation process will be discussed in relation to material and formulation characteristics to provide some general guidelines for process development.

Liquid-to-solid ratio (L/S ratio) has been found in the study of various formulations to be the most significant factor that affects the granule properties.[67,68] L/S ratio changes the friction force between the mixture and the inner wall of the barrel. A higher L/S ratio generates more agglomeration with less fracture due to a higher fill level and increased torque.[78] As a result, larger and denser granules can be produced.

The screw speed has an impact on the fill level. With a higher fill level in the barrel, higher friction force as well as higher torque are often observed. Increased screw speed and lower L/S ratio led to a shorter mean residence time and narrower residence time distribution. At high screw speeds, material flow inside the extruder is more plug flow as indicated by the narrow residence time distribution. Tu et al. studied a microcrystalline cellulose/polyethylene glycol 6000 formulation and found larger and more uniform granules were achieved with addition of more binder, a higher screw speed, and sufficient mixing.[78] However, changing the filling degree by varying the screw speed and feed rate did not significantly impact the PSD of granules with a lactose formulation.[70] In studying lactose granulation, Vercruysse found that the liquid addition method, pump orientation, pump type including peristaltic and piston pumps, tubing configuration (independent vs split tubing), nozzle diameters, liquid addition zones, and barrel-filling degree had no impact on mixing efficiency. More kneading zones and higher liquid feed rates were the only two factors that affected the mixing efficiency.

Nozzle design can influence the liquid dosing accuracy and consistency. Normally, the binder solution is injected right before the kneading block, which is also the place where pressure can build up due to the significantly increased fill level. The pressure

buildup can easily cause nuzzle clogging. To overcome this problem, a standard liquid nozzle can be modified to improve uniformity of the liquid feed rate. As demonstrated in Fig. 31.30, a tapered-tip nozzle design reduced the dead volume and the amount of material in front of the nozzle.[125] The Teflon capillary insert reduced the diameter nozzle and increased the liquid injection pressure. These designs reduced the clogging and led to a more consistent liquid feed. As discussed previously, the droplet size can affect the granule size. Addition of liquid to ungranulated powders in the dripping mode typically gives the broadest size distribution with un-granulated fines and a high percentage of lumps compared with addition using the spraying mode. Changing the nozzle diameter can change the droplet size. However, in the study of the lactose granulation by Vercruysse, the change in the nozzle diameter did not impact the moisture uniformity or the PSD of the granule.[70]

Binders can also be added in the form of foam. Thompson et al. successfully fed the foam binder using a side stuffer when the normal liquid injection device could not be used (see Fig. 31.31).[123] In foam granulation, the binder foam can rapidly spread over a large area of the powder during wetting; therefore, a lower amount of binder solution is needed compared to liquid injection. The granule properties from the foam granulation are comparable with the granules made using liquid injection.[123,124]

Different granulation behaviors have also been compared using dual-port granulation liquid injection and single-port granulation liquid injection. Dual-port granulation liquid injection facilitated the distribution of the granulation liquid and resulted in a lower and more stable torque during granulation. Changing the proportion of injection at the first injection port and second injection port did not have much impact on the torque.[79] Further improvement in moisture uniformity was not achieved by adding liquid injection zones.[70]

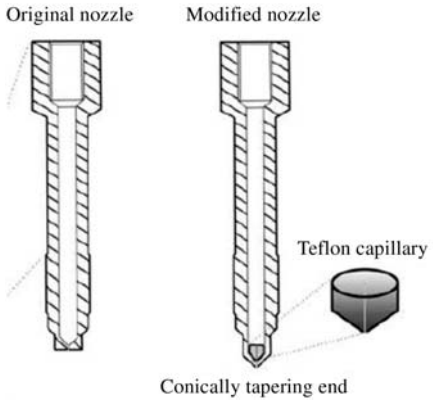


FIGURE 31.30 Liquid injection nozzle.[125]

Different binder concentrations in the binder solution also impact the granulation. Binder viscosity increases with increasing concentration. Although increased torque can be associated with higher binder viscosity during the TSE granulation process, the use of binder solution with higher viscosity can reduce fraction of the fines and, thus, facilitate generation of granules with unimodal distribution of particle size. In addition, it can also result in increased granulation strength and mean residence time and wider residence time distribution.[67] Furthermore, the binder is more effective when it is added as predissolved binder solution.[70]

In most cases, wet granulation via TSE is carried out at room temperature. At a higher barrel temperature, the powder mixture may dissolve faster in the granulation liquid compared to room temperature and lead to different granulation behavior. Vercruysse et al. reported that higher barrel temperatures led to fewer fines and stronger granules in theophylline/lactose/PVP formulation. The tablets made with granules granulated at a higher temperature had lower friability.[80]

Besides the process parameters and screw element design, attention should also be paid to the extruder geometry during process scale-up or transfer from one type of extruder to the other. Djuric et al. (2009) compared an APV Baker extruder with a Leistritz Micro extruder. A higher densification level was achieved for the extruder with a lower free chamber volume (Leistritz) in the granulation zone, which resulted in coarser granules and lower amount of fines. At the same screw speed, a higher tip speed was obtained for the extruder (Leistritz) with a larger screw diameter, which led to higher mechanical stress, resulting in coarser and stronger granules during process scale-up. Froude number, which takes into account the centrifugal force to the gravitational force, was used to determine the dynamic similarity of both extruders. The Leistritz extruder has a

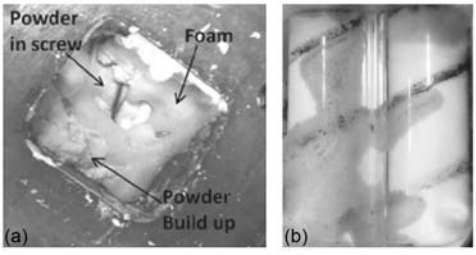


FIGURE 31.31 Different foam binder addition configurations. Image (a) shows the foam suspended above the moving screws when added directly to the open zone and powder buildup occurring on the barrel zone walls. Image (b) shows successful incorporation of the foam from a side-stuffer (as seen from a transparent viewing port above the screws while the foam entered out of view from the left-hand side). The direction of conveying for lactose was toward the top of the view in image (b).[123]

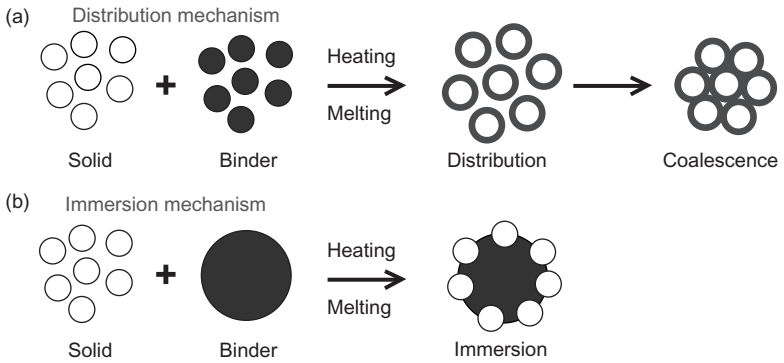


FIGURE 31.32 Agglomerate formation mechanisms in melt agglomeration. (a) Distribution mechanism; (b) immersion mechanism.[126]

higher Froude number, indicating higher centrifugal forces during granulation. When the free chamber volume is high, the feed rate can be increased without significant impact on the granule particle size.[76]

31.4.2 Continuous melt granulation

In the melt granulation, enlargement of the particles is achieved through the binding of solid particles with a molten binder. The binder used in the melt-granulation process is normally in a solid state at room temperature for ease of handling and melts or softens at elevated temperatures during the process. For continuous-melt granulation using a twin-screw extruder, the binder can either be preblended with other ingredients or subsequently fed into the extruder in a powder form, or it can be pre-melted and added continuously during granulation through a liquid injection nozzle. Melt granulation is more suitable for moisture-sensitive compounds and also has the benefit of avoiding a drying unit operation after granulation, which can increase the manufacturing efficiency.

In traditional melt granulation, binders are limited to a few choices with a melting point between 55°C and 85°C. There are generally two types of meltable binders: hydrophilic and hydrophobic. The former is often used in preparing IR dosage forms and is often limited to polyethylene glycols and poloxamers. The latter is mainly utilized in manufacturing ER products and may also serve as the rate-controlling materials. Examples include glyceryl behenate, carnauba wax, stearic acid, and glycerol monostearate. In continuous granulation using a twin-screw extruder, the more efficient mixing, heat transfer, and significantly less residence time allows a higher granulation temperature compared to a high-shear mixer. Therefore, a much lower amount and wider range of binders can be used, especially for producing IR products. In addition to the binders listed above, polymers such as polyvinylpyrrolidone (PVP), poly lactide-*co*-glycolide (PLGA), polyvinyl alcohol, and hydroxypropyl cellulose (HPC) can all be potentially used as the binder to facilitate granulation.[127,128] In certain cases, a thermally stable API with a relatively low melting point (eg, <120°C) can also serve as a binder for TSE process.

In the melt granulation process, the mechanism of granule nucleation and granule growth has been described as either immersion and layering or distribution and coalescence, as illustrated in Fig. 31.32.[81,126] When melted binder particles are smaller than powder particles or the viscosity of the binder is low, the liquid binder spreads over the surface of primary particles, and the wetted solids coalesce to form larger nuclei. When the melted binder particles are larger than the surrounding powder particles or the viscosity of the binder is high, the powder particles adhere to the surface of the droplet of the liquid binder. The granule grows as one layer of particles adheres to another layer of particles when the binder immerses out to the surface.[134] There is a strong correspondence between binder particle size and final granule size in the case of immersion and layering.

Mu and Thompson (2012) studied the mechanism of melt granulation using a twin-screw extruder.[81] In the study, two different grades of polyethylene glycol (PEG) were used as the binder, representing a binder with high and low viscosity. PEG was preblended with lactose and fed into the extruder in solid form. Inside the extruder, the heat transfer is mainly through convection and conduction. Because thermal conductivity of PEG is low, lactose powder would first be captured onto the softened PEG particle (as in the immersion model) rather than liquid PEG particle spreading around the lactose particle (as the distribution model). In a later stage, when the PEG particle is fully melted, its distribution to the outer surface of the particle by spreading would be more readily, especially in the mixing zone; therefore, granules could

grow through both mechanisms. The data indicated a strong correlation between PEG particle size with the granule particle size, which led the authors to conclude immersion mechanism was dominating. No study has been performed to study the mechanism when the binder is pre-melted and fed through a liquid injection nozzle. In this case, it is very likely that distribution and coalescence mechanisms will dominate the granulate formation and growth.

Continuous-melt granulation has been used to enhance the tableting properties of poorly compactible high-dose drugs, [82] reduce dose dumping for controlled-release drugs,[83] and reduce the tablet size for high-dose, modified-release formulation [84]. Lakshman et al. (2011) studied the tablet properties of a tablet containing 90% metformin HCl and 10% hydroxypropyl cellulose (HPC). Three granulation methods, melt granulation using a twin-screw extruder, and wet and solvent granulation using a high-shear granulator, were used in the study, and the results were compared. Due to the high water solubility of metformin HCl (about 1000 mg/mL), a small amount of moisture can dissolve a significant amount of the drug, which leads to poor powder flow, sticking, and picking during compression. Compared to the tablets made from wet granulation and solvent granulation, the tablets made from the continuous melt granulation process have higher hardness and lower friability when compressed under same compression conditions. In addition, the process is less sensitive to humidity changes in the manufacturing facility or moisture content in the granule compared to the wet and solvent granulation process. More detailed discussion of this study is available in Section 31.6.2. Keen et al. (2015) utilized a twin-screw extruder to make tramadol HCl extended-release formulation using glyceryl behenate as the binder, which is resistant to alcohol-induced dose dumping.[83] Different barrel temperature profiles, throughput rates, and lipid binder sizes were explored in the study and were found to have an impact on granule particle size. Vasanthavada et al. used the twin-screw extruder melt granulation to develop an imatinib mesylate modified-release formulation with a drug load of approximately 90%. In this study, melt granulation was compared with wet granulation. Tablets from the melt granulation process demonstrated higher tablet hardness, better friability, and slower drug release compared to the tablets prepared by a wet granulation process. The process robustness was also evaluated. None of the parameters, feed rate, process temperature, and screw speed, were found to significantly impact the final blend particle size and the drug release, demonstrating a robust granulation process.

31.4.3 Scale-up/scale-down considerations

Compared to a batch granulation process, the continuous granulation process has the benefit of avoiding scale-up by simply extending the granulation time and or operating on multiple extruders. Continuous granulation using a twin-screw extruder reaches steady-state quickly, normally after the duration of few residence times. After reaching the steady-state, granules can be consistently manufactured for days to meet the supply demand required. Keleb et al. reported successful continuous manufacturing of 8 hours of a hydrochlorothiazide /lactose/PVP formulation.[129] No significant differences in granule and tablet properties were observed over the entire granulation period.

In cases where the pilot-scale extruder cannot meet the commercial supply demand for a high-volume product, a scale-up to a commercial-scale extruder is needed. Scale-up of the extrusion process is introduced in Section 31.5. These scale-up methods are readily applicable to continuous granulation using a twin-screw extruder. Geometric similarity should be maintained between the different scales of the extruders. Most of the continuous wet/melt granulation processes are still operated under starve-fed conditions; therefore, a scale-up approach based on power or heat transfer may be more suitable compared with a volumetric scale-up method. For both continuous wet and melt granulation, the shear energy generated is much lower compared to a HME process due to different screw configurations and the absence of a flow-restricting die block. The power assumption is normally lower than 10%. Therefore, API degradation due to heat/shear energy is less of a concern compared to a HME process.

Unlike a batch granulation process in which the materials are preweighed and discharged into a batch, the solid powder and liquid binders are continuously fed into the extruder for granulation in a continuous granulation system. The feeding accuracy and consistency directly impact the final composition of the product, which makes the control of the powder and liquid feeding system critical during scale-up and scale-down. One major objective of granulation is to improve the API flow properties; therefore, the API in a granulation formulation may have poor flow and high variability in physical properties. This makes the feeding challenging, especially when the API is directly fed into the extruder without preblending with other excipients. Cartwright et al. studied different loss-in-weight (LIW) feeding systems for a poorly flowing API.[130] Three different screw designs (see Fig. 31.33) of K-tron feeders were studied for the K-tron rigid frame design. Fine concave screws and

FIGURE 31.33 Transport screws used in K-Tron LIW feeder.[130]

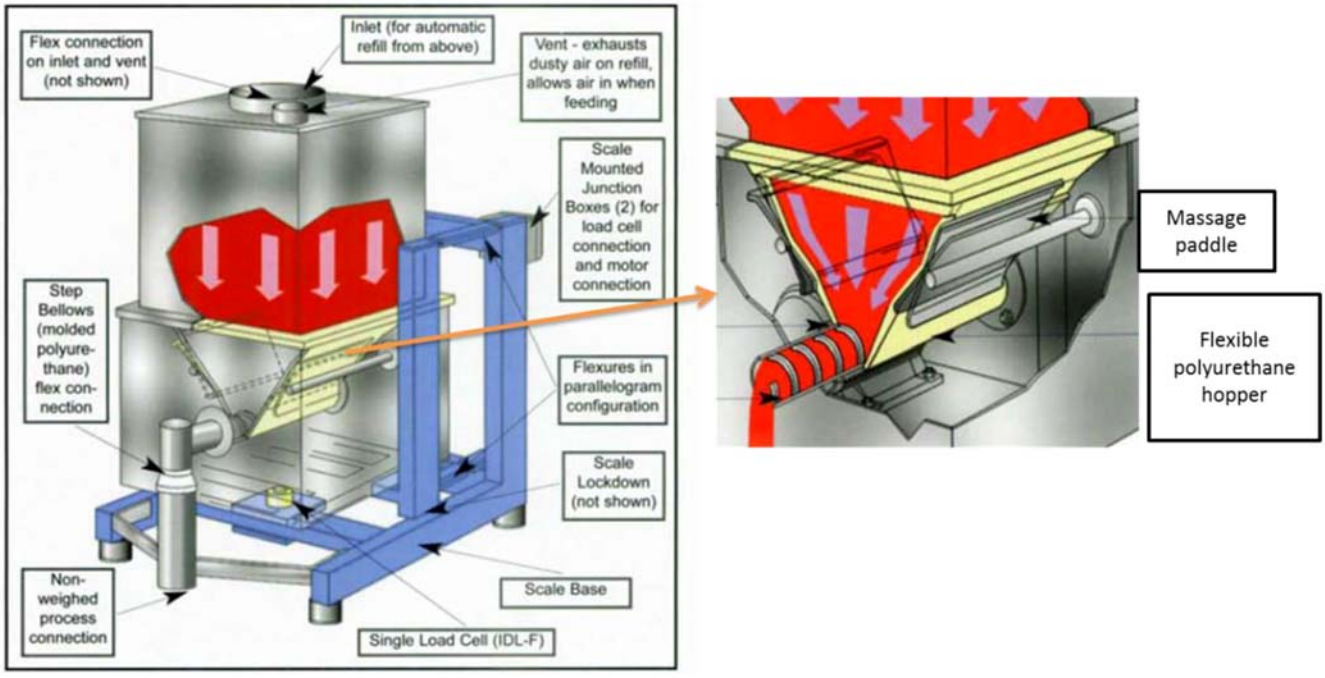


FIGURE 31.34 Brabender flexwall feeder.[131]

flighted auger screws could be used for free-flowing and dense materials. However, it caused powder compaction within the screw and between the screw and the feeder barrel wall for low bulk density and poor flow material. A core and spiral screw conveyed the poor flow material more effectively than the fine concave and flighted auger screw, but the feed speed was limited. In addition, frequent powder bridging was observed in the powder hopper due to the scimitar blade. Material compaction against the hopper walls and the widest arc of the scimitar blade was also noticed. The flexible frame loss in weight feeder (see Fig. 31.34) has hopper walls made from polyurethane or silica and massage paddle to agitate powder through. It can eliminate bridging and avoid powder compaction. Cartwright et al. used the flexible frame loss-in-weight feeder and achieved more consistent flow compared to the K-tron feeders.[130]

31.4.4 Modeling of continuous granulation

Ideally, a simulation model of the twin-screw granulator can provide insights on improving process knowledge and performing qualitative study, thus reducing the amount of experiments needed. For example, DEM and CFD models can provide particle tracking information, which can be used to construct

the residence time distribution along the spatial domain for a continuous corotating screw granulator.

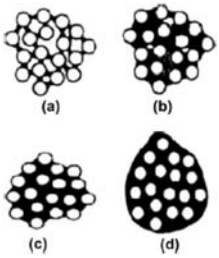
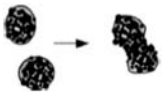
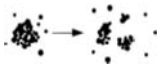
However, first principle simulations of the flow in corotating intermeshing twin-screw extruders are highly challenging due to the complex geometry of the rotating screw and free surface flow in partially filled conveyed sections. On the other hand, the calibration and validation of simulation models are crucial steps in assessing their value in granulation process modeling. Typically, adjustments or tuning in model parameters through calibration are necessary to improve model accuracy and obtain reliable models. However, many of these parameters are difficult to measure in the field, creating additional challenge for simulation. Most of the modeling approaches to date were carried out using default parameter values or best-guessed values to describe the rate of nucleation, growth, and breakage. Some modeling techniques are computationally challenging. In the case of the discrete element method (DEM), it can be very computationally costly to simulate a large number of particles with small integration time steps.

A comprehensive review by Kumar et al. on applying population balance and discrete element modeling to high-shear wet granulation (HSWG) provides a good foundation for modeling of a continuous granulation process.[85] Bouffard et al. and Barrasso et al. implemented a bidirectional coupling of population balance model (PBM) and DEM using a commercially available software.[86,87] To avoid the challenges in the DEM approach, Eitzlmayr et al. have employed the smoothed particle hydrodynamics (SPH) method on a Newtonian, temperature-independent flow, taking advantages of SPH's meshless nature and the inherent capability to simulate free surface flows.[88]

To develop a microscale model such as DEM, the incorporation of microscale material properties such as wet granulate yield strength, Young's modulus, and asperity size are required. Accurately measuring these properties are important but challenging in some cases, especially the granule yield strength at different wetness levels. Table 31.7 shows the interdependencies of particle scale parameters and system parameters.[85] The combination of PBM (system level) and DEM (particle level) is the most promising approach for modeling of the continuous granulation process. These models at different scales should be linked using multiscale integration framework. This approach has been summarized by Kumar.[85]

So far, the model validation has been qualitative and focuses on model fitting of the endpoint of

TABLE 31.7 Size Changing Mechanisms Occurring in HSWG[85]

Mechanism	Particle formation	Characteristics	Equations
Wetting and nucleation		**(a)** Pendular—looks like bridge, but particles not immersed in liquid **(b)** Funicular—thicker bridges but not completely filled **(c)** Capillary—particles at edge of cluster not completely wetted by liquid **(d)** Droplet—all particles completely wet	$\tau_{wetting} = \frac{2V_0^2}{\pi^2\varepsilon_s^2 r_d^4 R_{pone}} \frac{\mu}{\gamma_{IV}\cos\theta}$ $\frac{\partial n}{\partial t} = B^0\delta(l - l_0)$
Growth-aggregation		• When dealing with systems that exhibit aggregation, it is more convenient to use particle volume rather than particle size, since volume is conserved • The success of collisions, ie, aggregation can be a function of particle size, liquid content and powder properties and operational factors such as bed height, powder velocity and shear	$B_{agg}(x) = \frac{1}{2}\int_0^x \beta(x-y,y)n(x-y,t)n(y,t)dy$ $D_{agg}(x) = n(x,t)\int_0^\infty \beta(x,y)\,n(y,t)dy$
Growth-layering		• It is more convenient to use particle size as the internal coordinate when describing layering or growth	$\frac{\partial n}{\partial t} = \frac{\partial}{\partial x}\left[n\frac{dx}{dt}\right](x,t)$
Breakage		• The complexity of breakage models extends from binary breakage models to full particle distributions represented by breakage and selection functions or empirical models	$\frac{\partial n(u)}{\partial t} = \int_0^\infty K_{break}(y, x-y)\xi_{break}(y)n(y,t)dy - \xi_{break}(x)n(x,t)$

granulation outcomes, such as granulate size distribution and their physical properties. Barrasso et al. have demonstrated the parameter estimation and validation for a multidimensional PBM used for a twin-screw granulation process.[89] As the measurements techniques for endpoint determination progress, more vigorous and complex models can be calibrated and validated utilizing the available experimental data.

31.5 PROCESS SCALE-UP

To consistently produce the extrudate with a predefined quality, it is essential to evaluate and understand both the formulation and the manufacturing process using a systematic approach. To scale-up a manufacturing process, a good starting point can be an Ishikawa (Fishbone) diagram, which provides a comprehensive overview of all variables involved. Fig. 31.35 is an example fishbone diagram for the HME process.[24] As discussed in the previous sections, a unique aspect of the extrusion process is that it consists of several processing zones. As shown in Table 31.5, HME can typically divided into the: (1) feeding zone, (2) solid conveying zone, (3) melting zone, (4) mixing zone, (5) melt conveying zone, and (6) the discharging/shaping zone with an optional devolatilization/venting zone. Thus, in successfully scaling up an extrusion process, the requirements for individual zones and the overall process need to be considered simultaneously.

31.5.1 Extrusion process scale-up

The ultimate goal of extrusion process scale-up is to generate a large quantity of extrudate while maintaining critical quality attributes. As illustrated in Fig. 31.15, the extrudate quality is determined by scale independent key system parameters, including specific energy, residence time distribution, product temperature, and potentially pressure. Therefore, those parameters should be maintained constant during scaling. In the most straightforward and simple scaling, the extrusion batch size increases by increasing the extrusion time or duplicating the same extruder used in the small scale. Under such a scenario, the extrusion process remains the same, and the process parameters should be kept unchanged. However, when scaling requires an equipment change, the scaling could become challenging as the key system parameters are not directly adjustable and, in some cases, even difficult to measure accurately. A number of publications proposed scale-up theories of HME. Nakatani reported a successful extrusion scaling by introducing an adiabatic index aimed at keeping the melt temperature constant.[90] The suitability of a one-dimensional computational extrusion simulation was demonstrated for process scale-up and transfer.[91,92] In this section, a few classical scale-up strategies will be discussed. Although these strategies provide a starting point of the scaling process, further refinement and optimization of extrudate parameters (free volume, screw configuration, die geometry) and process parameters (screw speed, feed rate, and barrel temperature)

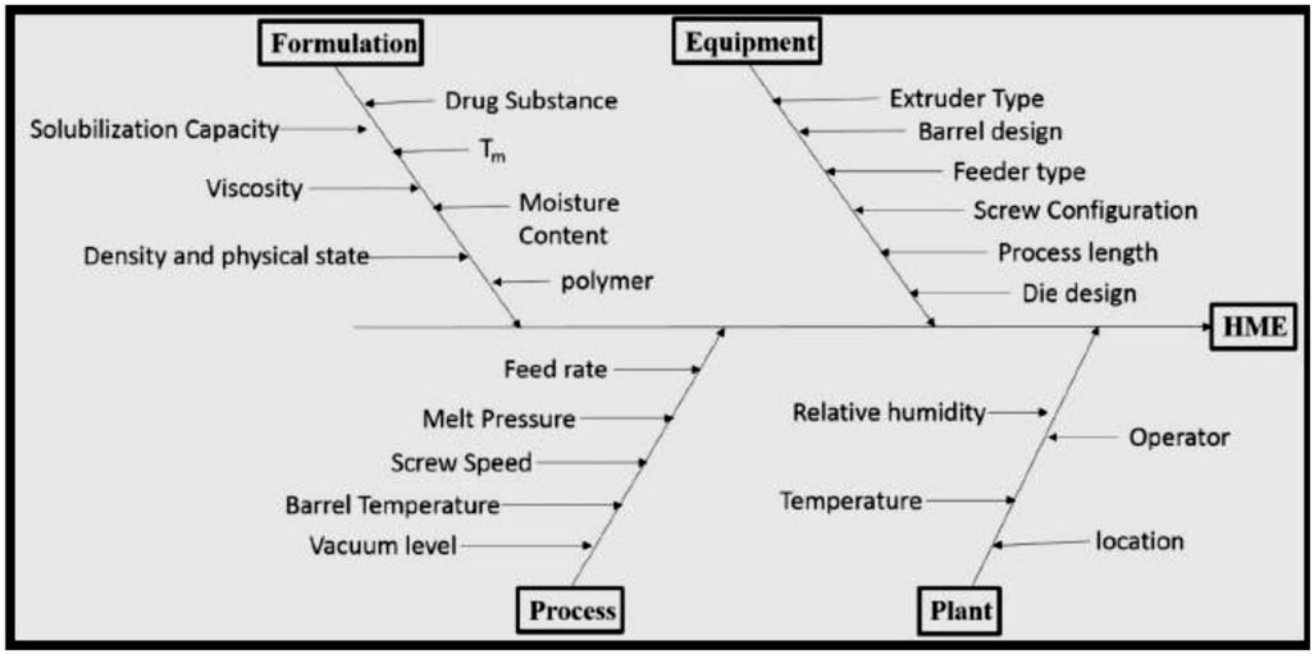


FIGURE 31.35 Ishikawa diagram of hot-melt extrusion process.[24]

are often required to identify the proper operating conditions.

31.5.1.1 Geometric similarity

Similar to scaling for all other unit operations, geometric similarity between extruders should be maintained during scaling. It is recommended to adhere to the same design of extruder and ideally remain with the same equipment manufacturer. To achieving geometric similarity, the following three extruder and screw dimensional characteristics should be examined and kept as constant is possible:

1. Length/diameter (L/D): The ratio between screw/barrel length and screw diameters (L/D) and the length of each processing zone are important for accomplishing desired processing tasks such as conveying, melting, and mixing. This ratio, to a great extent, is determined by the extruder design.
2. Screw diameter ratio (D_o/D_i): The ratio between D_i and D_o impacts free volume, shearing rate, and mixing capability. Most twin-screw extruders today are manufactured with diameter ratios between 1.45 and 1.75.
3. Screw design: The screw plays an important role in determining product quality, with conveying and mixing elements configured in a specific sequence to accomplish the desired task in each processing zone by controlling the residence time distribution and energy input. The sections of a screw configuration and their relative positions can be scaled using extruder outer diameter D_o as a scaling factor.[93] For conveying sections, both pitch and section length are calculated as a ratio to D_o, and those ratios should be maintained constant between scales. The same method applies to scaling of the kneading section length (see Fig. 31.36). However, the number of disks and offset angles should be kept constant if possible. Since an exact scale-up of each screw element for all machine sizes may not be available, the screw configuration and zone position may be changed, and, as a result, the adjustment of process parameters would likely be necessary.

31.5.1.2 Classic scale-up strategies

The classic scale-up strategies focus on the process limiting factors, including feed capacity (volumetric scale-up), motor power (power scale-up), and heat input (heat transfer scale-up).

31.5.1.2.1 Volumetric scale-up

Volumetric scale-up of the extrusion process focuses on maintaining a constant mean residence time. The volumetric scale-up strategy applies when free volume limits the throughput, that is, the extruder is operating at its volumetric limits. This occurs when the feed zone is full as a result of low material density and/or conveying capacity; increasing the screw speed increases the throughput.

When two extruders are considered geometrically similar, the degree of fill and average shear rate are also expected to be reproduced. For an extruder with the same screw diameter ratio, the same shear rate is expected when operating at the same screw tip speed. The anticipated scale-up volumetric throughput is believed to follow a cubic law and can be estimated according to Eq. (31.11).[132]

$$Q_2 = Q_1 \times \left(\frac{D_2}{D_1}\right)^3 \tag{31.11}$$

Where Q_2 is anticipated throughput of extruder with diameter D_2; Q_1 is measured throughput of the extrudate with diameter D_1.

However, when the screw diameter ratio between extruders is different, but similar screw geometries with similar screw tip speed are used, Eq. (31.11) can be modified as Eq. (31.12):

$$Q_2 = Q_1 \times \left(\frac{FreeVolume_2}{FreeVolume_1}\right) \tag{31.12}$$

Eq. (31.11) is also modified to accommodate the situation in which the screw rotation speed, expressed as N, is different between scales (Eq. 31.13).[132]

$$Q_2 = Q_1 \times \left(\frac{D_2}{D_1}\right)^3 \times \frac{N_2}{N_1} \tag{31.13}$$

As the residence time depicts mass flow patterns in an extruder and plays a critical role in determining product characteristics, it is recommended to obtain this information, particularly in the case of scaling-up and transferring processes to a different extruder with differing geometries. Comparing to other key system parameters, the residence time distribution is fairly easy to measure using a pulse stimulus response technique [94–96]. Fig. 31.37 shows a typical set-up for residence time distribution measurement.[97] During the measurement, a tracer is added into the feeding section at $T_{0,}$ and the concentration of the tracer is measured at the exit of extruder over the time. The concentration of the tracer can be measured using either an offline or online method (eg, NIR or Raman).

31.5.1.2.2 Power scale-up

Power scale-up is based on constant SME, which means constant energy per unit mass. Maintaining constant SME during scaling is critical, particularly for the manufacture of ASD systems, because the energy input is essential to achieve the desired product

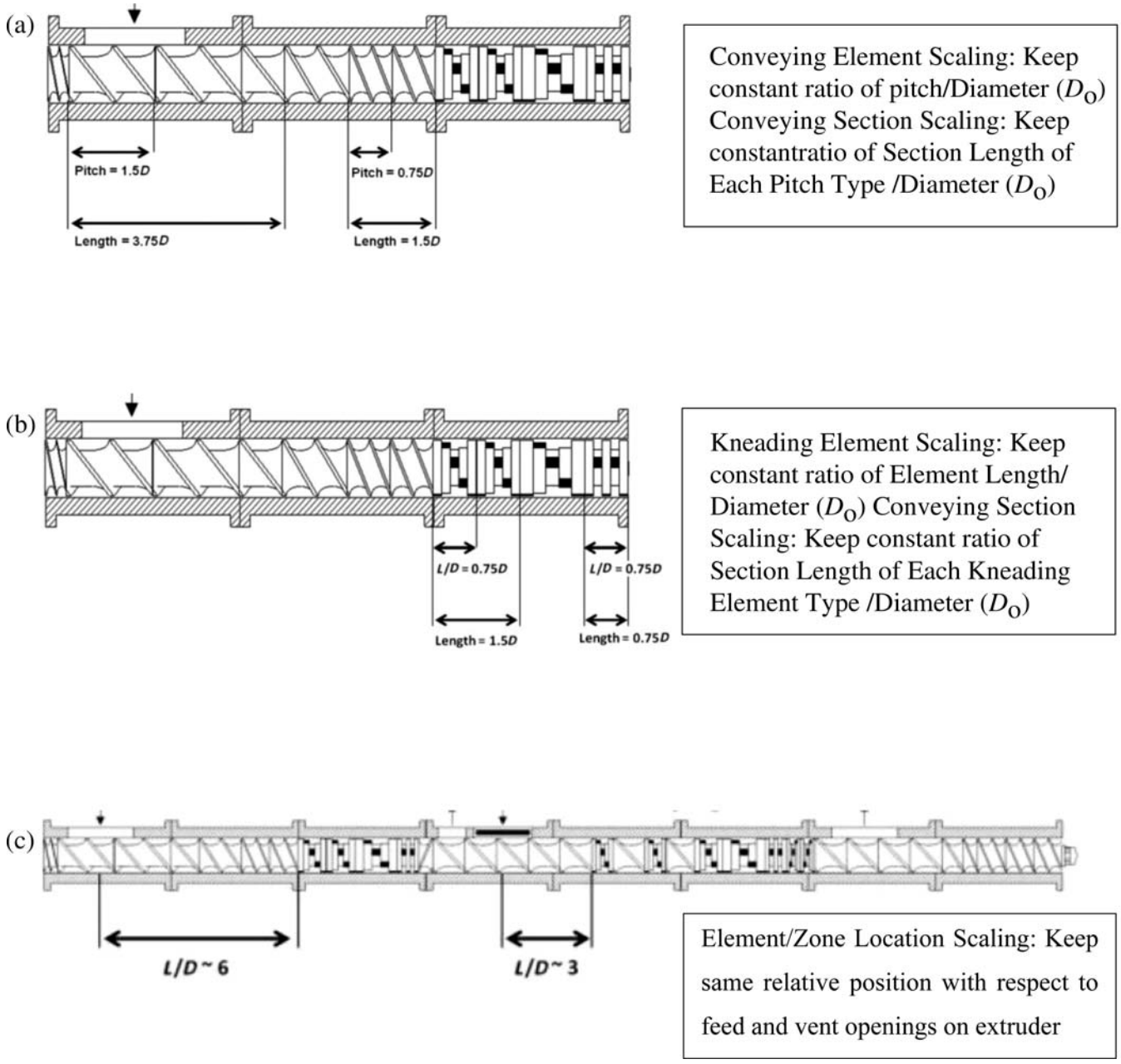


FIGURE 31.36 Scale-up of screw design. (a) Conveying section scale-up; (b) kneading section scale-up; (c) screw element/zone location scale-up. Modified from Ref. [93].

attributes when processing in the solubilization regime. SE is the sum of SME and specific thermal energy (STE). STE input is governed by heat conduction from barrel to material. This section will focus on the scale-up strategy based on SME.

One of the most straightforward SME calculations is shown in Eq. (31.6). The premise of applying this calculation during scaling is similar screw geometry, same percentage of fill, and equivalent screw speed as measured by tip velocity. The maximum throughput can be calculated with a given SME for the small extruder and available power for the large extruder. When comparing this throughput to the one calculated using the volumetric scale-up strategy, a disparity may be observed. This could be due to the torque density difference among extruders. Torque density (see Eq. 31.14) describes the relationship between power and free volume. When scaling to an extruder with lower torque density using the power scale-up strategy without considering torque density, the

FIGURE 31.37 Residence time distribution measurement setup.[97]

calculated throughput is less than the one obtained through the volumetric scale-up strategy. This may result in the process operating at lower degree of fill, higher shear rate, and longer residence time. Thus, the torque density should be assessed prior to the application of the power scale-up strategy.

$$Torque\ density = \frac{Torque}{Centerline\ Distance^3} \tag{31.14}$$

SME is a function of torque, screw speed, and throughput in which screw speed and throughput are adjustable process parameters, while torque is a response variable. Thus, this strategy can also be used to evaluate the relationship between power (as function of torque and screw speed) and screw speed. In general, if an increase in throughput is desired, screw speed should be increased congruent to maintaining constant SME.

31.5.1.2.3 Heat transfer scale-up

The heat transfer scale-up strategy can be adopted when the extrusion process is limited by heat transfer, and the desired melt temperature may not be achieved. Thermal heat transfer in TSE is dependent on the degree of fill, barrel surface area, temperature gradient between product and barrel, and residence time. Direct scale-up for heat transfer is based on the heat transfer area, that is, a square law (Eq. 31.15). Based on this strategy, the heat transfer coefficients should be constant among scales to maintain the same melt temperature. As with other scale-up assumptions, the degree of fill and screw diameter ratio must be the same.

$$Q_2 = Q_1 \times \left(\frac{D_2}{D_1}\right)^2 \tag{31.15}$$

For extruders with different screw diameter ratios, Eq. (31.15) should be modified based on the true value of the inner barrel surface area.

A heat transfer limit is not common in pharmaceutical applications, as product temperature is mainly dependent on the viscous dissipation generated by converting SME rather than barrel heating or cooling. In general, the same barrel temperature profile can be employed initially during scaling. However, due to the circular geometry of the extruder inner barrel surface, we should realize that increasing the extruder size results in the decreased thermal heat transfer per unit mass. Thus, when scaling with this strategy, a large extruder could be operated at lower degree of fill as compared to the small-scale machine due to increased melt temperature. Lower fill at a constant screw speed also increases the average shear rate and residence time. Scaling based on heat transfer limitation is a complex process; therefore, caution should be taken when using this strategy. Additional experiments and modeling might be required to yield proper scaling results.

31.5.1.2.4 Simulation-assisted scale-up strategies

Extrusion simulation allows for process scale-up with high confidence and in a cost-efficient manner. Simulation-assisted scale-up is particularly useful when geometric similarity and the requirement for classic scale-up strategies cannot be met. The ultimate goal of simulation tools is to produce a glass extruder so that the process is fully understood, and all properties are visible. Kohlgrüber summarized the possibilities and limitations of extrusion modeling.[98] Based on the number of spatial model dimensions, modeling approaches can be categorized into: (1) zero-dimensional models (0-D), which focus on the balance of the complete extruder or an individual section; (2) one-dimensional models (1-D), which focus on the process along the extruder axis; and (3) two- or three-dimensional models (2-D, 3-D), which consider the cross-section and require defined the boundary condition to better reflect the process.[98]

The 3-D models, typically based on complete numerical solutions of motion equations using FEA or similar techniques, provide an accurate and detailed picture of the flow field during the extrusion process. Polyflow is a commercially available 3-D simulation software. This model assumes a fully filled extruder under nonisothermal conditions and can track the simulated trajectories of particles. The software output is a 3-D description of flow in an extruder, including distributions of temperature, pressure, shear rate, and other responses. Additionally, the software can quantify dispersive and distributive mixing as well as predict stagnation areas. However, since this model can only simulate the viscous fluid-filled zone, it is not possible to model an entire extrusion process along the length of the screw.[99]

Alternatively, 1-D models are typically based on lumped-parameter methods, in which the equations of motion are integrated throughout the cross-section. By trading off the resolution, 1-D models offer the desired information within a shorter period of time with

simpler inputs requirement (eg, material properties not understood at the molecular level are sufficient). The commercially available 1-D models are summarized in several published articles.[99,100] Akro-Co-Twin Screw is one of the earliest marketed models, while Ludovic is the most widely used in the pharmaceutical industry.[99,100] Although different 1-D models can be developed based different theories and/or calculation methods, the common simulation outputs include temperature, pressure, fill ratio, viscosity, shear rate, energy consumption, and residence time distribution. The sample outputs from Sigma and Ludovic are shown Fig. 31.38 and Fig. 31.39.[98,101] The emphasis on the change along the screw axially for particularly difficult to measure system parameters (ie, energy distribution, residence time distribution, and material temperature) enables thorough process understanding and easy process scale-up. It was reported that computational extrusion process simulation using Loduvi could identify high-energy intake locations and allow optimization of screw design to ensure scalability.[102] It is also demonstrated that the WinTXS software package can be tuned to match simulation results with experimental data by using proper inputs, including material properties and process information. A correlation can be established between product quality and simulation output, which assisted an appropriate scaling.[92]

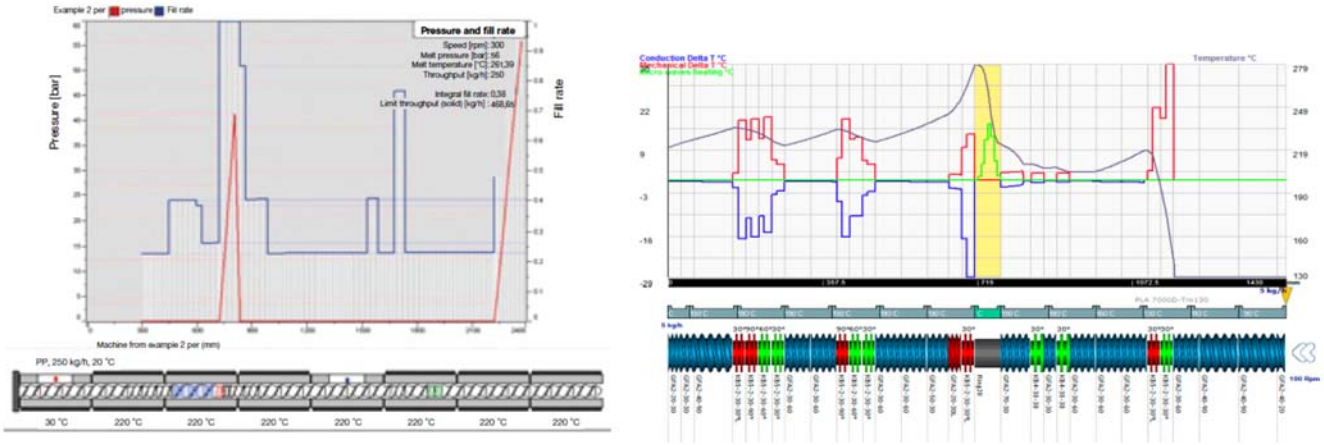


FIGURE 31.38 Sample output using sigma and Lodovic program.[98]

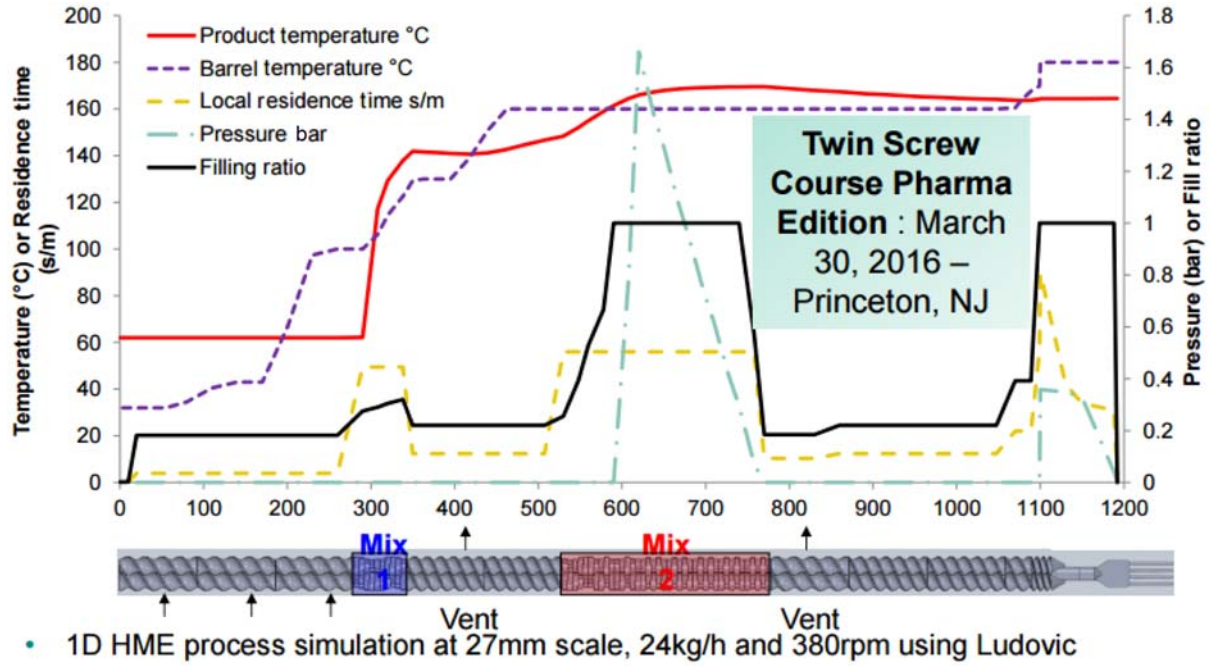


FIGURE 31.39 Modeling of HME via SCC's Ludovic.[101]

31.5.2 Product and process understanding, control strategy, and PAT

Quality by design (QbD) promotes a thorough understanding of the product and manufacturing process via a systematic approach. The developmental activities as per QbD expectations as shown in Fig. 31.40 are part of the quality risk-management process. In the context of extrusion, once the quality target product profile (QTPP) linked directly to safety and efficacy has been identified, the development focus should follow a series of structured risk assessments. A comprehensive formulation and process evaluation based on critical quality attributes (CQAs) should be performed at the initial risk assessment. The results from this risk assessment help prioritize the development activities. For instance, minimizing SME input while maintaining sufficient mixing is the study focus when developing an ASD with a material prone to thermal degradation. However, when developing an extrusion granulation process with a similar material, understanding the impact of degree of fill and residence time distribution might take precedent over the efforts of reducing SME input. Design space, as one of the important elements of a QbD approach, offers operational flexibility. The expectation for design space is that the process understanding should be achieved at a multivariate level. That is, the interdependence between the product quality and individual process parameters and the additive influence from multiple parameters are clearly identified and understood. As the extrudate quality is directly driven by key system parameters, as shown in Fig. 31.15, the deployment of those scale-independent process responses is an ideal way to bridge the gap between quality attributes and process independent parameters. Developing the design space around the key system parameters provides manufacturing flexibility and seamless scale-up. As the sources of variability and their impacts are understood, a control strategy span to raw material, in process control (IPC), and process design must be developed to ensure CQAs are consistently met. Multiple control strategies for the same process are possible, as long as they can be justified through a rigorous quality risk-management process.

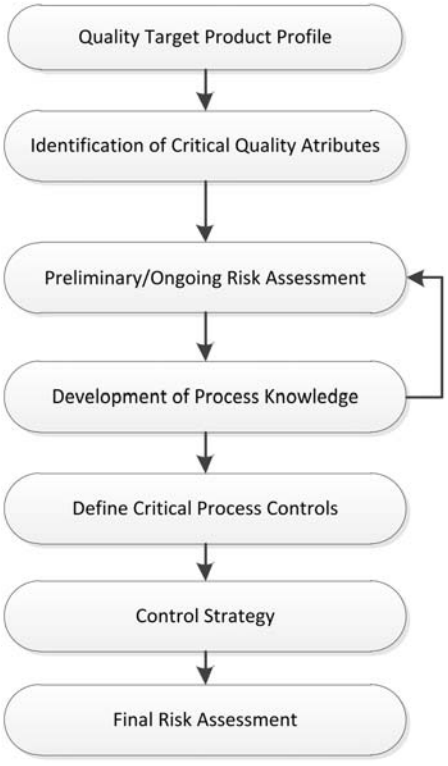


FIGURE 31.40 Product and process development following QbD approach.

One of the other important components of QbD is process analytical technology (PAT). PAT encompasses online, in-line, or at-line measurement for enhanced process understanding and improved control for product quality. The implementation of PAT yields many manufacturing advantages. Two specific advantages to the extrusion process are noted here: (1) PAT enables continuous manufacturing of quality extrudate by monitoring and identifying deviations in real-time,[103] and (2) PAT improves manufacturing efficiency by eliminating time consuming offline analysis, and this is particularly true for hot-melt polymer matrix formulation as the extraction of drug substance from the matrix can be complicated.

PAT tools have long been used by the plastics and food industries for monitoring, analyzing, and characterizing the extrusion process and product in-line. The schematic of PAT integration into the melt extrusion process is illustrated in Fig. 31.41.[104] Raman and near-infrared (NIR) are two widely studied spectroscopy-based analytical techniques for in-line or online monitoring of the melt extrusion process.[105–107] Saerens et al. reported using Raman to monitor the API solid state change from crystalline to amorphous. When combined with principal component analysis (PCA) and partial least squares analysis (PLA), they also demonstrated that Raman is able to simultaneously quantify the API concentrations in process samples.[103] The suitability of Raman for PAT application was also presented by Tumuluri et al. for clotrimazole and ketoprofen formulations.[107] Similar to Raman spectroscopy, NIR can provide both quantitative and qualitative information on melt and/or extrudate. It can be used to monitor the content uniformity and verify the extrusion performance in terms of residence time distribution. Islam et al.

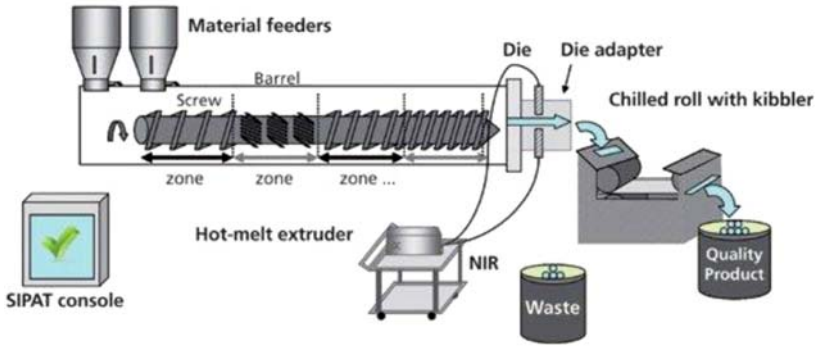


FIGURE 31.41 Schematic of PAT integration into melt extrusion.[104]

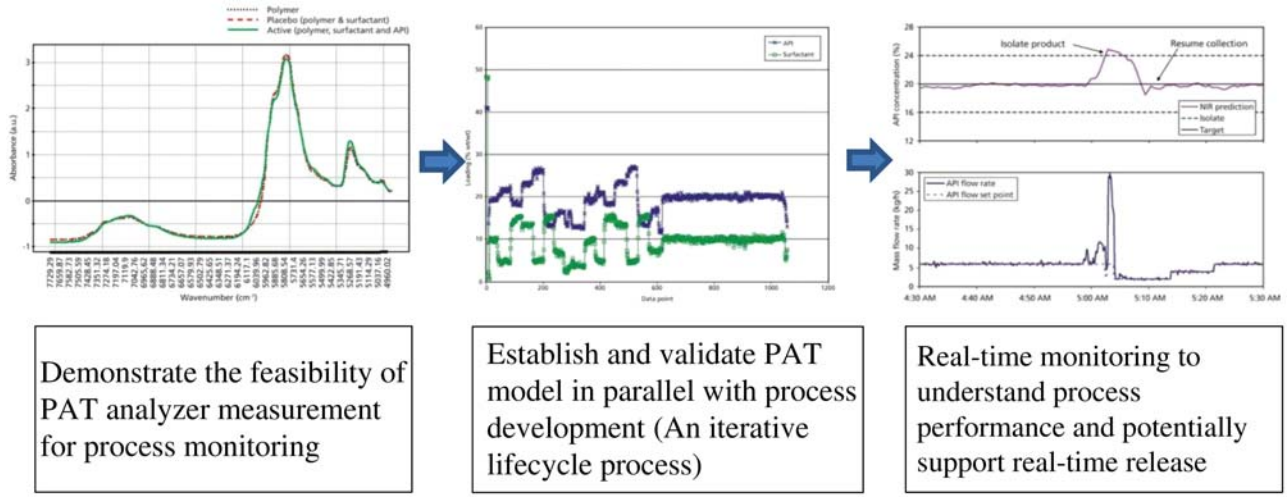


FIGURE 31.42 Establish PAT model for process monitoring. Modified from Ref. [104].

utilized in-line NIR spectroscopy to assess the effect of the process critical parameters and identify the critical quality attributes (CQA).[108] Saerens et al. suggest that in-line NIR spectroscopy could be a potential PAT tool for monitoring polymer–drug solid-state behavior while determining API concentration.[108] Additionally, other tools, including ultrasound and UV–vis spectroscopy have also been reported in the literature for in-line characterization of extrusion processes.[109–111]

Implementing PAT for real-time quality assurance requires a validated model to correlate the CQAs with process analyzer measurements. Mathematical transformation is typically needed when process analytical data is in the format of spectra or chemical images. Validating a PAT model is an iterative process that can last throughout the product lifecycle. The process map of establishing a PAT model for melt extrusion process monitoring is shown in Fig. 31.42.[104,133] Once the feasibility of a PAT analyzer measurement for the formulation is demonstrated, developing a calibration model is initiated. The model is then challenged and validated in parallel with development and ready to be implemented for routine commercial production once proven it is robust to all foreseen variability. A properly developed PAT model should be able to monitor the process performance based on the CQAs of interest and quickly identify any deviations from predetermined acceptance criteria. If a problem is identified, the suspect extrudate can be diverted into waste, and the product collection will resume when the process fault is resolved. Thus, the real-time quality assurance is achieved with PAT, that is, a deviation or process fault can be detected immediately upon occurring and lead to accurately isolating this product portion without impacting the entire batch. In addition, PAT eases the demand between process monitoring and manufacturing support by simplifying the multivariate nature of the process into a single data stream representing product quality.

31.6 CASE STUDIES

Over the last decade, many new products have been developed utilizing the TSE process to achieve the desired drug product profile. This section provides several examples of product developmental using TSE to enable or improve process, product quality and performance.

31.6.1 Improving oral absorption of BCS class II drugs

Utilizing HME to manufacture ASD formulations to improve oral absorption of insoluble drugs offers many advantages when feasible. Various fundamental aspects of product development are addressed in the relevant chapters of this book. Two examples discussed in this section illustrate the importance of understanding API, product, and process when using HME.

Rezulin is the first commercial product containing ASD manufactured using HME process. Troglitazone, the active component of Rezulin, is a BCS class II drug with low and highly variable bioavailability. At the time of the product approval in the 1990s, Park-Davis/Warner-Lambert had already developed an organic solvent-based spray-drying process to prepare a povidone-based ASD of troglitazone with enhanced bioavailability. Forecasting a large commercial volume for Rezulin, concerns arose about the cost, environment, health, and safety impact associated with the anticipated increasing use of large quantity of organic solvents during production. A concerted effort to identify an alternative solvent-free manufacturing process led to the HME process using a twin-screw extruder. The TSE process initially developed using a Leistritz 18 mm intermeshing corotating twin-screw extruder was successfully scaled up to a Leistritz 50 mm extruder for the commercial production. The extrusion processing time is less than 1 minute for the batch size of 5000 kg.

Noxafil delayed-release tablets contains a melt-extruded ASD of posaconazole, a weakly basic triazole antifungal drug. Posaconazole exhibits pH-dependent solubility. It dissolves in acid conditions but rapidly precipitates in the upper small intestine. The first generation Noxafil, a suspension of crystalline form, exhibits highly variable bioavailability with a substantial food effect. The melt-extruded formulation was compared to the suspension at a 100 mg dose in a crossover study in healthy volunteers examining fed-state and fasted-state plasma profiles. The results, presented in Fig. 31.43, show substantial oral bioavailability improvements with reduced food effect.[112] However, Posaconazole undergoes significant degradation at temperatures above its melting point of 170°C, presenting a challenge for commercial scale extrusion. To overcome this problem, a formulation was designed in which the molten polymer (HPMCAS) was used to solubilize posaconazole below its melting point. This allows for production of the material at reduced temperatures to improve the stability of the drug.

31.6.2 Improving processing characteristics through modification of API surface

Metformin is a biguanide antihyperglycemic agent most widely prescribed for the treatment of type 2 diabetes. Its hydrochloride salt is freely soluble in water

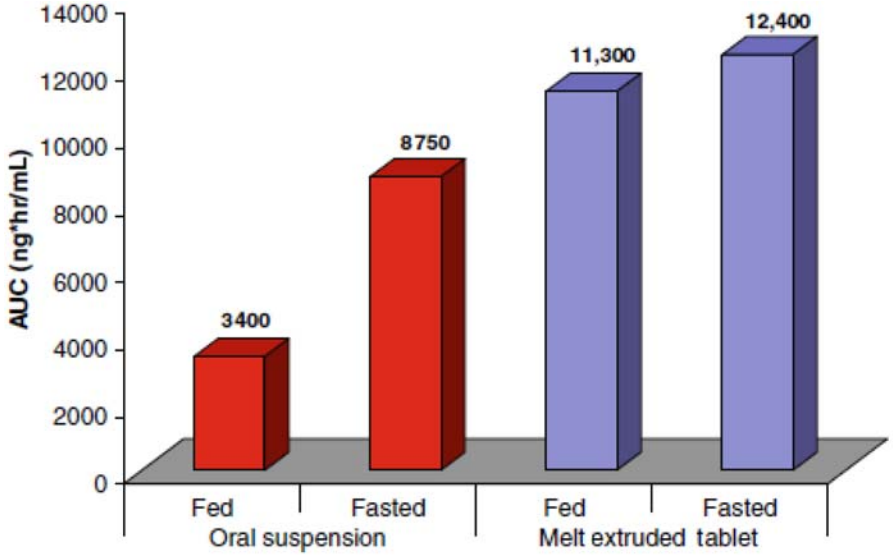


FIGURE 31.43 Posaconazole oral bioavailability: suspension containing crystalline drug vs tablets containing amorphous solid dispersion.

with a pK_a of 12.4 and melting point of 223–226°C. Due to its high daily dose up to a total of 2000 mg, the common strengths of metformin tablets include 500 mg, 850 mg, or 1000 mg, which indicate that (1) high drug loading is required, especially when developed as a fixed-dose combination or ER dosage form and (2) formulation processing behaviors will be dominated by the API properties. Because of high solubility (about 300 mg/mL)[113] metformin HCl dissolved in moisture can form solid bridges upon moisture desorption.[82] In worst cases, extensive formation of solid bridges could transform free-flowing powder into a solid block overnight. As a result, variable bulk and tablet properties of metformin HCl made using wet granulation have been observed, including granulation flowability, density, compactability, and the hardness, disintegration, and dissolution of the tablet. To overcome these problems, Lakshman et al. (2011) investigated melt granulation (MG) of metformin HCl with hydroxypropyl cellulose (HPC) as the meltable binder using a 16 mm Thermo Prism twin-screw extruder to poorly compactible high-dose API. A high (90%) drug-load tablet formulation containing 1025 mg of metformin was melt granulated at a temperature above the glass transition of HPC but below the melting point of metformin HCl (see Table 31.8). It was found that good compressibility can be achieved using MG process depending on process conditions (Fig. 31.44). When compared to the aqueous and solvent wet granulation (WG) processes, significantly improved hardness and friability results were obtained for tablets made with the MG process. Tablet hardness of MG tablets was also found to be less sensitive to changes in moisture content during compression and subsequent storage. The low-moisture contents of the granules (about 0.2% w/w) is believed to be conducive to averting moisture-induced physical transformations in metformin HCl tablets, thus avoiding change of tablet hardness with time, variable powder flow behavior, and problems with dissolution stability. In addition, the study also shows that the TSE granulation process can decrease the need for relatively large amounts of excipients often required for overcoming processing challenges posed by API physicochemical and mechanical properties, thus reducing tablet sizes of high-dose drugs.

The significantly enhanced compactibility of metformin HCl tablets can be attributed to the effective modification of the API surface properties through highly efficient and more consistent coating of the binder (HPC) through the melt-extrusion process. This is indicated by the rank-order improvement in tablet bonding among the three processes studied by the authors.[82] Based on the principle of the granulation process, the coating efficiency by HPC is expected to be lower with aqueous-based WG process when compared with solvent-based WG process (ethanol/isopropanol, 95:5, v/v) because both the API and HPC are dissolved in the granulation fluid during the mixing process, while solubility of the API in alcohol is much lower. Since the API is not miscible with the molten HPC, the highest coating efficiency is achieved with the TSE process, resulting in most effectual modification of the API surface properties. This is consistent with the study findings in our laboratory using a poorly compressible API as a model drug to manufacture extended-release hydrophobic matrix tablets. A powder blend consisting of about 90% API and about 6% rate-controlling polymer and other excipients were granulation using an 18 mm twin-screw extruder equipped with six barrels and with the die block removed from the end barrel (Coperion ZSK laboratory extruder). Solid materials were fed into the extruder at a feed rate of 1 kg/hour

TABLE 31.8 Melt-Granulation Process Parameters Evaluated for Metformin Tablets[82]

Process parameter	Set point	Variation ± set point or fixed
HPC level	9%	5–15%
Feed rate	30–35 g/min	± 10 g/min
Screw speed	210 rpm	± 90 rpm
Extrusion temperature	160°C	140–180°C

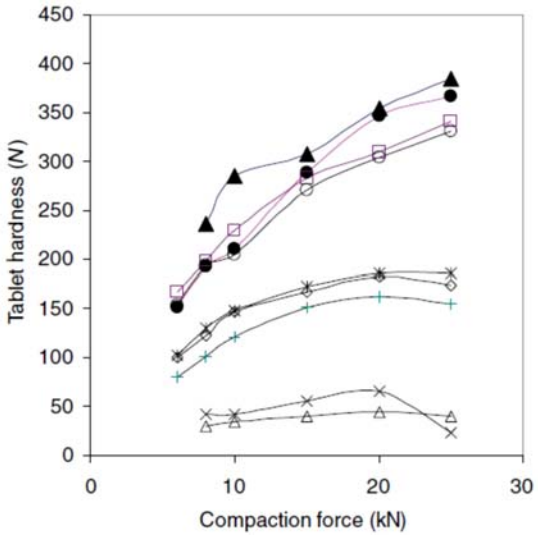


FIGURE 31.44 Compaction profile of metformin Hcl melt granulations.[82] The feed rate, screw speed, highest processing zone temperature, and magnesium stearate level were, respectively: •: 40 g/min, 120 rpm, 140°C and 0.75%; +: 20 g/min, 300 rpm, 180°C and 0.75%; ○: 40 g/min 120 rpm, 140°C and 1.25%; □: 40 g/min, 120 rpm, 180°C and 1.25%; *: 30 g/min, 210 rpm, 160°C and 1.00%; ×: 20 g/min, 300 rpm, 140°C and 1.25%; ◇: 30 g/min, 210 rpm, 160°C and 1.00%; ▲: 40 g/min, 120 rpm, 180°C and 0.75%; and △: 20 g/min, 300 rpm, 140°C and 0.75%.

using a K-Tron twin-screw volumetric feeder followed by granulation using conveying elements at a screw speed of 120 rpm and a barrel temperature of 150°C. Because the melting point of the API (770°C) is significantly higher than the glass transition temperature (T_g) of the polymer (41°C) and because there is no API-polymer miscibility, the surface of the API is modified effectively by the polymer through the extrusion process, enabling tablet compression. Tablets of the same formulation were also prepared with a solvent-based high-shear WG process using polymer solution in ethanol as the granulation fluid followed by drying, sieving, and compression. Because the API is insoluble in ethanol, the granulation-drying process also facilitated similarly effective coating of the API surface. As a result, tablets made using both processes exhibited similar compactability and dissolution. Vasanthavada et al. (2011) also studied the high-dose ER formulations of imatinib mesylate with a drug load of 89% by melt granulation using a 16 mm Thermo Prism twin-screw extruder (ie, a 1074-mg tablet containing 956 mg API).[84] The process temperature of 185°C is below the melting point of the API (212°C) but above the T_g values of polymers, which included HPC, ethyl cellulose (EC), and hydroxypropyl methylcellulose (HPMC). A factorial design of experiment (DOE) designed to investigate the effects of process variables, such as feed rate, process temperature, and screw speed, over a wide range demonstrated robust performance with respect to granulation and tablet characteristics as well as drug release. Acceptable compressibility and tablet hardness can be obtained using as little as 5% polymer for melt granulation. In addition, tablets produced using the MG process exhibited higher hardness and lower friability when compared with those made with the WG process, consistent with the findings from MG study of metformin tablets previously discussed. Evaluation of the tablets using confocal Raman microscopy revealed that the API remained as unmelted crystals, and polymers were finely distributed (coated) on the surface and in between API particles, confirming the importance of effective coating of the polymer in improving API processibility by modifying its surface properties.

31.6.3 Continuous granulation for manufacturing extended-release tablets using a twin-screw extruder

Clarithromycin is a macrolide antibiotic used to treat many different types of bacterial infections. It is a weak base with pKa = 8.7 and pH-dependent solubility. Biaxin XL is the extended-release hydrophilic matrix tablets of clarithromycin. It was a high-volume product manufactured using wet granulation in a high-shear granulator followed by fluid-bed drying, milling, compression, and film coating. In Zu et al.,[114] Biaxin XL tablet is used as a model drug product to investigate the feasibility of continuous wet granulation using a twin-screw extruder.

Granulation was performed by feeding the preblended formulation of the commercial tablets[115] consisting of 50% API, 20% HPMC, and other excipients into an 18 mm corotating twin-screw extruder (Coperion ZSK laboratory extruder) equipped with six barrels without a die. Granulation fluid (water) was added through liquid injection into a designated barrel. Process variables studied over a wide range include liquid feed rate, screw configuration and speed, liquid-to-solid ratio, and position of liquid injection using an experimental design provided in Table 31.9.

TABLE 31.9 Study Design of Granulation Process Parameters for Clarithromycin Extended-Release Tablets Using TSE

Run	Screw configuration	Liquid feed rate (kg/h)	Screw speed (rpm)	Injection position	L/S ratio
1	#1	0.090	120	3–1	0.090
2	#1	0.132	120	3–1	0.132
3	#1	0.282	120	3–1	0.282
4	#1	0.432	120	3–1	0.432
5	#1	0.582	120	3–1	0.582
6	#1	0.282	200	3–1	0.282
7	#1	0.282	300	3–1	0.282
8	#1	0.282	120	4–1	0.282
9	#1	0.282	120	4–2	0.282
10	#2	0.282	120	3–1	0.282
11	#2	0.282	200	3–1	0.282
12	#2	0.282	300	3–1	0.282
13	#2	0.090	300	3–1	0.090
14	#3	0.282	120	3–1	0.282
15	#3	0.282	200	3–1	0.282
16	#3	0.282	300	3–1	0.282
17	#4	0.282	120	3–1	0.282
18	#4	0.282	200	3–1	0.282
19	#4	0.282	300	3–1	0.282
20	#5	0.090	120	3–1	0.090
21	#5	0.132	120	3–1	0.132
22	#5	0.282	120	3–1	0.282
23	#5	0.282	120	3–2	0.282
24	#5	0.282	120	4–1	0.282
25	#5	0.282	300	3–1	0.282

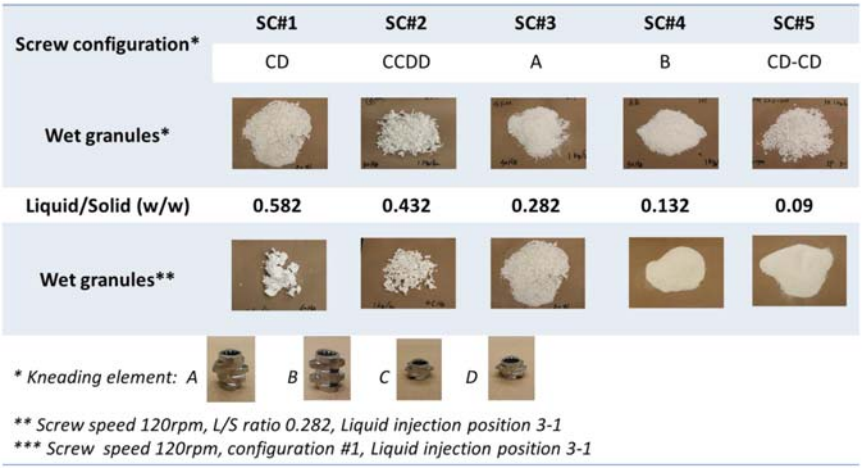


FIGURE 31.45 Influence of screw configuration and percentage liquid on granulation characteristics of clarithromycin ER tablets.

The influence of these process parameters on characteristics of the resulting granules (particle size, strength, and density), milled granules (PSD, strength, and density) and tablets (solid fraction, hardness, and dissolution) was evaluated. Among the process parameters investigated, liquid-to-solid ratio, screw speed, and screw configuration showed the most significant effect on the granule properties (Fig. 31.45). For example, increasing the percentage of liquid or the liquid-to-solid ratio results in larger and wetter granules and higher granule strength. Granules are smaller when a less aggressive kneading zone is used. The granule size decreases with an increase of screw speed. However, these large differences in granulation properties can be mostly equalized by controlling the milling process of the dried granulation (eg, using a 840 μm band and Fitzmill with impact forward, 3000 rpm). The only factor that affects the milled granule particle size is the liquid-to-solid ratio. Other process parameters including screw configuration do not have a significant impact on the milled granule properties. Tablet hardness, porosity, and dissolution are also comparable within a wide range of percentage of liquid of 9−28.2% when compared with the tablets manufactured using high-shear granulation. Increased porosity and faster dissolution were observed for tablets made with >43% liquid, using commercial tablets as a reference, though tablet strength remains essentially unchanged. When compared to the high-shear granulation process, TSE not only renders significantly more homogeneous admixture of the formulation composition, as indicated by the nearly uniform potency values across different granulation sieve cuts, but it also requires a considerably reduced amount of liquid to achieve effective granulation. For example, approximately 35% of liquid is typically added using the high-shear granulator, while liquid levels as low as 9% have been shown to be sufficient to produce tablets with the same quality when using the TSE process. As a result, a drastically reduced drying cycle is expected due to a lower moisture content and homogeneity of the resulting wet granules.

To evaluate the applicability of the study findings to other products, Depakote ER tablet was investigated as a second-model drug product using a similar but reduced study design by focusing on the effects of liquid-to-solid ratio while keeping the other process parameters constant. Depakote ER is also an extended-release tablet based on a hydrophilic matrix system that provides approximately 20 hours of in-vivo absorption[116]. It was also a high-volume product manufactured using wet granulation in a high-shear granulator followed by fluid-bed drying, milling, compression, and film coating. The active ingredient, divalproex sodium, is a stable and permeable compound with an acidic pKa of 4.8. It is freely soluble in the intestinal pH where most of the dose is released. In contrast to Biaxin XL, Depakote ER tablets contain high loading of high-solubility API and high level of high-viscosity-grade HPMC. Nevertheless, the study results were found to be similar to those of Biaxin XL. For instance, similar dissolution was observed with tablets manufactured using percentage of liquid ranging from 9% to 23% using the high-shear granulated product (about 14% liquid) as a reference. However, accelerated dissolution was observed with tablets made from considerably overgranulated and much larger granulation, which was milled using a 1000-μm band and Fitzmill with impact forward, 3000 rpm (see Fig. 31.46). It should be noted that an increase in the dissolution rate does not take place until approximately 40−50% of drug release. Hence, the change can be most likely be attributed to the gradual loss of matrix integrity during dissolution, because large particle size combined with low porosity of the granules made

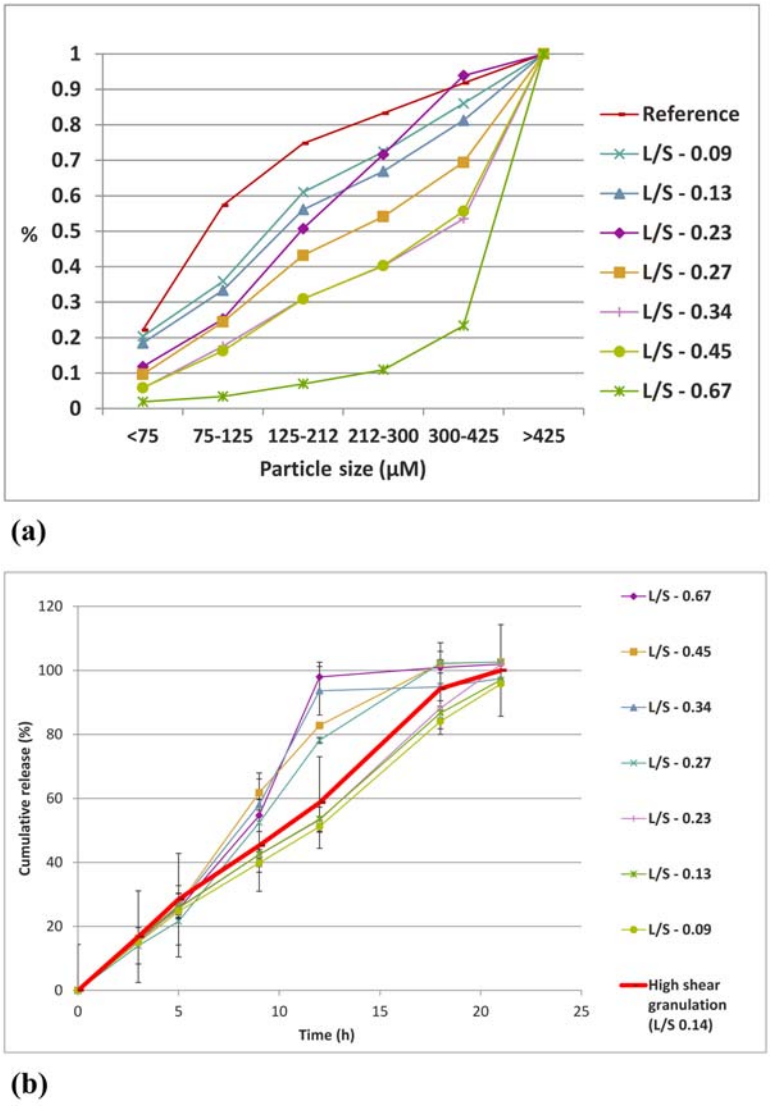


FIGURE 31.46 Influence of percentage liquid on (a) particle size distribution and (b) dissolution of divalproex sodium ER tablets.

with high-percentage liquid can alter water penetration and gel structure of the matrix. It is found that the impact is more pronounced with the ER formulation of divalproex sodium containing high-viscosity-grade HPMC when compared with clarithromycin ER formulation that contains low-viscosity-grade polymer, since the former usually results in larger particle size and higher density. In principle, this type of dissolution problem can be resolved by controlling PSD through modifying the milling condition.

31.6.4 Extended-release opioid tablets with abuse-deterrent properties

The abuse of prescription opioid drug products has been a growing problem in the United States. One of the successful formulation strategies to deter this abuse is to prepare a dosage form with high physical strength to prevent crush and extraction. Both Opana ER extended-release oxymorphone hydrochloride tablets and Nucynta ER extended-release tapentadol hydrochloride tablets are manufactured using INTACT, a proprietary technology developed at Grünenthal GmbH. Poly(ethylene oxide) is used as the drug-release retardant and matrix former. The manufacturing process consists of blending, extrusion, cooling, cutting, forming, and coating. During the melt extrusion, the poly(ethylene oxide) based matrix formulation is melted and pressurized. As a result, the melt-extruded matrix possesses strong physical strength in comparison to the tablets prepared using a conventional tableting process. The resulting PEO matrix tablet exhibits a breaking strength of greater than 500 N. As shown in Fig. 31.47, the melt-extruded formulation could not be crushed by a mortar and pestle.[117]

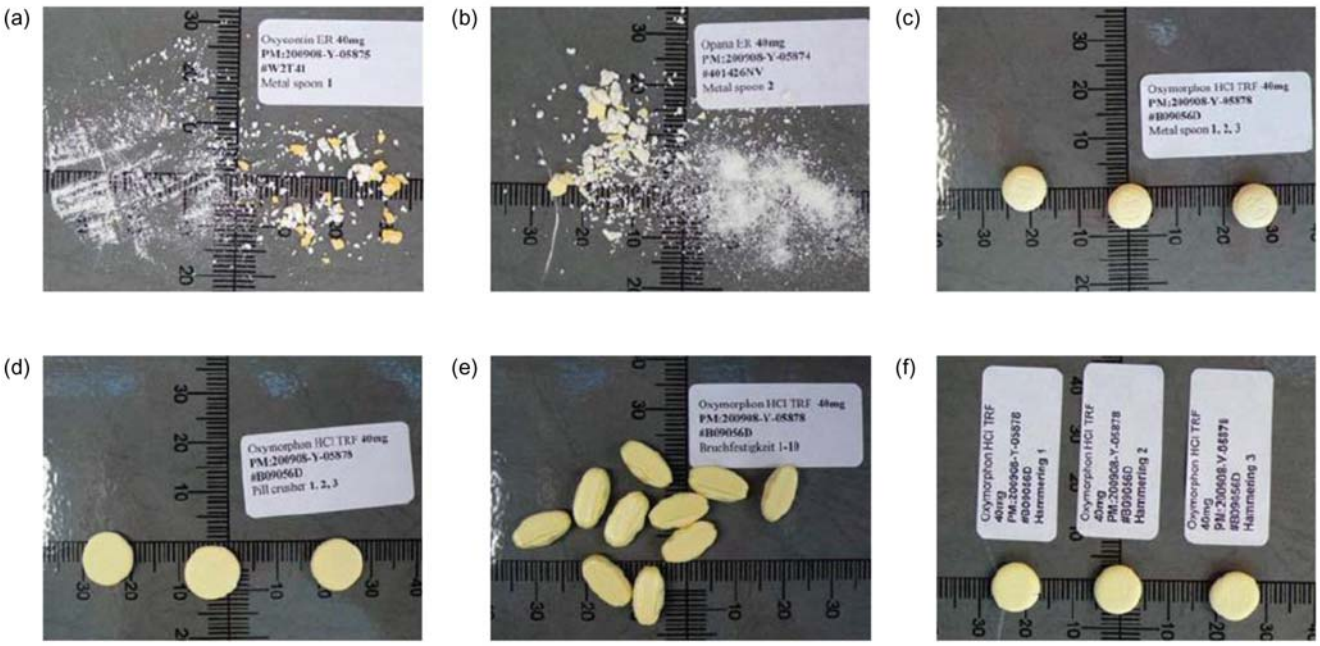


FIGURE 31.47 High mechanical strength of Polyox matrix tablets prepared using melt extrusion.[117]The effect of compression with two spoons on (a) traditional extended-release oxycodone hydrochloride tablets without tamper resistance, (b) traditional extended-release oxymorphone hydrochloride tablets without tamper resistance, and (c) crush-resistant oxymorphone hydrochloride extended-release tablets prepared using melt extrusion. Effect on TSE extended-release tablets by (d) professional pill crusher, (e) tablet hardness tester, (f) hammer apparatus.

31.7 SUMMARY

The rapid adoption, implementation, and expanded applications of the extrusion technology in the pharmaceutical industry resemble what occurred in the food and plastics industry five decades ago. In the early 1960s, the food and plastics industries began to abandon batch mixers and embrace extrusion technology in order to more efficiently produce products of more consistent quality and at a lower cost. The batch mixer is a large-volume discontinuous mixer. In comparison, a twin-screw extruder is a small-volume continuous mixer. Valuable knowledge, technical know-how, and manufacturing experience gained in these industries have allowed expeditious adoption and implementation of this established processing technology to enable formulation development of challenging drug molecules and efficient pharmaceutical manufacturing.

TSE is being adopted by the pharmaceutical industry at very rapid pace. Pharmaceutical companies are not only actively utilizing the TSE process for manufacturing novel drug-delivery systems but are also replacing less efficient and less robust batch processes with TSE extrusion. At the same time, twin-screw extruder suppliers have downsized and redesigned equipment for production in the GMP environments. Additional efforts have also been made to design TSE systems for testing early development materials available only in limited quantities. Excipient companies are engineering novel materials suitable for extrusion process. At the same time, regulatory agencies are encouraging the pharmaceutical industry to adapt flexible, robust, and efficient continuous manufacturing process in order to build product quality into process design.

References

1. Giles HF, Mount EM, Wagner JR. *Extrusion: the definitive processing guide and handbook*. Norwich, NY: William Andrew Inc; 2005.
2. Bruce C, Manning M. Melt extruded thin strips containing coated pharmaceutical actives, USPTO, 2012/0308635 A1, 2012.
3. Potluri R, Todd D, Gogos C. Mixing immiscible blends in an intermeshing counter-rotating twin screw extruder. *Adv Polymer Technol* 2006;**25**:81–9.
4. Thompson M, Sun J. Wet granulation in a twin-screw extruder: implications of screw design. *J Pharm Sci* 2010;**99**:2090–103.
5. Nakamichi K, Nakano T, Yasuura H, Izumi S, Kawashima Y. The role of the kneading paddle and the effects of screw revolution speed and water content on the preparation of solid dispersions using a twin-screw extruder. *Int J Pharm* 2002;**241**:203–11.
6. Crowley MM, Zhang F, Repka MA, Thumma S, Upadhye SB, Battu SK, et al. Pharmaceutical applications of hot-melt extrusion: Part I. *Drug Dev Ind Pharm* 2007;**33**:909–26.
7. Madan S, Madan S. Hot melt extrusion and its pharmaceutical applications. *Asian J Pharm Sci* 2012;**7**:123–33.
8. Particle Sciences. 2011. Hot melt extrusion technical brief. Elsevier. <http://ees.elsevier.com/ajps/default.asp>.

9. Eslami, H. Understanding screw design for film extrusion process. <http://www.macroeng.com/understanding-screw-design-for-film-extrusion-process.php>.
10. Todd DB. Introduction to compounding. In: Todd DB, editor. *Plastics compounding, equipment and processing*. Cincinnati: Hasner Gardner Pbulication; 1998. p. 1–12.
11. Zimmermann D. Screw design for co-rotating twin-screw extruders. *Plast Additiv Compound* 2004;**6**:38–41.
12. Anderson PG, Lechner F. *Co-rotating fully intermeshing twin-screw compounding: advancements for improved performance and productivity*. Orlando: ANTEC; 2012.
13. Paul EL, Atiemo-Obeng VA, Kresta SM. *Handbook of industrial mixing: science and practice*. Hoboken, NJ: Wiley; 2004.
14. Martin C. Twin screw extrusion for pharmaceutical processes. In: Repka AM, Langley N, DiNunzio J, editors. *Melt extrusion: materials, technology and drug product design*. New York, NY: Springer; 2013. p. 47–79.
15. Padmanabhan B. Understanding the extruder processing zone: the heart of a twin screw extruder. *Plast Additiv Compound* 2008;**10**:30–5.
16. Kessler T, Breitenbach J, Schmidt C, Degenhardt M, Rosenberg J, Krull H. Process for producing a solid dispersion of an active ingredient, USPTO, US2009/0302492 A1, 2009.
17. Cheremisinoff NP. *Guidebook to extrusion technology*. Englewood Cliffs, NJ: Prentice Hall; 1993.
18. Kleinebudde P, Lindner H. Experiments with an instrumented twin-screw extruder using a single-step granulation/extrusion process. *Int J Pharm* 1993;**94**:49–58.
19. Mollan M. Historical overview. In: Ghebre-Selassie I, Martin C, editors. *Pharmaceutical extrusion technology*. New York, NY: Marcel Dekker, Inc; 2003. p. 69–98.
20. Repka AM, McGinity JW, Zhang F, Koleng JJ. Hot-melt extrusion technology. In: Swarbrick J, Boylan J, editors. *Encyclopedia of pharmaceutical technology*. New York, NY: Marcel Dekker; 2002. p. 203–66.
21. Morott JT, Pimparade M, Park JB, Worley CP, Majumdar S, Lian Z, et al. The effects of screw configuration and polymeric carriers on hot-melt extruded taste-masked formulations incorporated into orally disintegrating tablets. *J Pharm Sci* 2015;**104**:124–34.
22. Jani R, Patel D. Hot melt extrusion: an industrially feasible approach for casting orodispersible film. *Asian J Pharm Sci* 2015;**10**:292–305.
23. Kostic M, Reifschneider L. Design of extrusion dies. In: Lee S, editor. *Encyclopedia of chemical processing*. New York, NY: Taylor & Francis; 2006. p. 633–49.
24. Patil H, Tiwari RV, Repka MA. Hot-melt extrusion: from theory to application in pharmaceutical formulation. *AAPS PharmSciTech* 2016;**17**:20–42.
25. Perdikoulias J, Dobbie T. Die design. In: Ghebre-Selassie I, Martin C, editors. *Pharmaceutical extrusion technology*. New York, NY: Marcel Dekker; 2003. p. 99–109.
26. Breitenbach J, Grabowski S, Rosenberg J. Extrusion von polymer-wirkstoff-gemischen zur herstellung von Arzneiformen. *Spektrum der Wissenschaft* 1995;**11**:18–19.
27. Claeys B, Vervaeck A, Hillewaere XKD, Possemiers S, Hansen L, De Beer T, et al. Thermoplastic polyurethanes for the manufacturing of highly dosed oral sustained release matrices via hot melt extrusion and injection molding. *Eur J Pharm Biopharm* 2015;**90**:44–52.
28. Pederson A. Erosion-based drug delivery. *Manuf Chem* 2006;**11**:1–6.
29. Quinten T, De Beer T, Vervaet C, Remon JP. Evaluation of injection moulding as a pharmaceutical technology to produce matrix tablets. *Eur J Pharm Biopharm* 2009;**71**:145–54.
30. Roth W, Burst A, Zietsch M. Abuse resistant melt extruded formulation having reduced alcohol interaction, EP20090779544, 2009.
31. DiNunzio JC, Miller DA. *Formulation development of amorphous solid dispersions prepared by melt extrusion. Melt extrusion*. New York, NY: Springer; 2013. p. 161–203.
32. Breitenbach J. Melt extrusion: from process to drug delivery technology. *Eur J Pharm Biopharm* 2002;**54**:10.
33. Bialleck S. *Herstellung von Polysaccharidpellets mittels Schmelzextrusion*. Rheinischen Friedrich-Wilhelms-Universität Bonn; 2011.
34. Roblegg E, Jäger E, Hodzic A, Koscher G, Mohr S, Zimmer A, et al. Development of sustained-release lipophilic calcium stearate pellets via hot melt extrusion. *Eur J Pharm Biopharm* 2011;**79**:635–45.
35. Treffer D, Wahl P, Markl D, Koscher G, Roblegg E, Khinast GJ. Hot-melt extrusion as a continuous pharmaceutical manufacturing process. In: Repka AM, Langley N, DiNunzio J, editors. *Melt extrusion: materials, technology and drug product design*. New York, NY: Springer; 2013. p. 363–96.
36. Meuser F, Gimmler N, Van Lengerich B. A systems analytical approach to extrusion. In: Ho C-TH, Karwe MV, Kokini JL, editors. *Food extrusion science and technology*. New York, NY: Marcel Dekker; 1992. p. 619–30.
37. Yacu WA. *Extruder selection, design, and operation for different food applications. Advances in food extrusion technology*. Boca Raton, FL: CRC Press, Taylor & Francis Group; 2011. p. 23–68.
38. Reiner R. General overview of the compounding process: tasks, selected applications, and process zones. In: Lemens K, editor. *Co-rotating twin-screw extruder*. Cincinnati: Carl Hanser Verlag GmbH & Co. KG; 2007. p. 57–89.
39. Gao Y, Muzzio FJ, Ierapetritou MG. A review of the residence time distribution (RTD) applications in solid unit operations. *Powder Technol* 2012;**228**:416–23.
40. Anderson PG. The Werner & Pfleiderer twin screw co-rotating extruder system. In: Todd DB, editor. *Plastics compounding, equipment and processing*. Cincinnati: Hasner Gardner Publications; 1998.
41. Lowinger M. MELT EXTRUSION-process development: scaling a melt extrusion process from conception to commercialization. *Am Pharm Rev* 2011;**14**:80.
42. Huber G. Twin screw extruders. In: Riaz MN, editor. *Extruders in food applications*. Baca Raton, FL: CRC Press; 2000. p. 211.
43. Henrist D, Remon JP. Influence of the process parameters on the characteristics of starch based hot stage extrudates. *Int J Pharm* 1999;**189**:7–17.
44. Verhoeven E, De Beer T, Van den Mooter G, Remon JP, Vervaet C. Influence of formulation and process parameters on the release characteristics of ethylcellulose sustained-release mini-matrices produced by hot-melt extrusion. *Eur J Pharm Biopharm* 2008;**69**:312–19.
45. Shibata Y, Fujii M, Sugamura Y. The preparation of a solid dispersion powder of indomethacin with crospovidone using a twin-screw extruder of kneader. *Int J Pharm* 2009;**365**:8.
46. McCrum NG, Buckley CB. *Principles of polymer engineering*. Oxford: Oxford University; 1997.
47. Rauwendaal C. *Polymer extrusion*. Los Altos Hills, CA: Carl Hanser Verlag GmbH & Company KG; 2014.
48. Verreck G, Decorte A, Heymans K, Adriaensen J, Cleeren D, Jacobs A, et al. The effect of pressurized carbon dioxide as a temporary plasticizer and foaming agent on the hot stage extrusion process and extrudate properties of solid dispersions of itraconazole with PVP-VA 64. *Eur J Pharm Sci* 2005;**26**:349–58.
49. Verreck G, Decorte A, Heymans K, Adriaensen J, Liu D, Tomasko D, et al. Hot stage extrusion of p-amino salicylic acid with EC using CO2 as a temporary plasticizer. *Int J Pharm* 2006;**327**:45–50.
50. Newman A. *Pharmaceutical amorphous solid dispersions*. New York, NY: Wiley; 2015.
51. Shah N, Sandhu H, Choi DS, Chokshi H, Malick AW. *Amorphous solid dispersions: theory and practice*. New York, NY: Springer; 2015.
52. LaFountaine JS, McGinity JW, Williams RO. Challenges and strategies in thermal processing of amorphous solid dispersions: a review. *AAPS PharmSciTech* 2015;**17**:43–55.

53. Marsac PJ, Li T, Taylor LS. Estimation of drug–polymer miscibility and solubility in amorphous solid dispersions using experimentally determined interaction parameters. *Pharm Res* 2009;**26**:139–51.
54. Marsac PJ, Shamblin SL, Taylor LS. Theoretical and practical approaches for prediction of drug-polymer miscibility and solubility. *Pharm Res* 2006;**23**:2417–26.
55. Tao J, Sun Y, Zhang G, Yu L. Solubility of small-molecule crystals in polymers: D-mannitol in PVP, indomethacin in PVP/VA, and nifedipine in PVP/VA. *Pharm Res* 2008;**26**:11.
56. Friesen D, Shanker R, Crew M, Smithey D. Hydroxypropyl methylcellulose based spray-dried dispersions: an overview. *Mol Pharm* 2008;**5**:1003–19.
57. Gupta SS, Meena A, Parikh T, Serajuddin AT. Investigation of thermal and viscoelastic properties of polymers relevant to hot melt extrusion, I: Polyvinylpyrrolidone and related polymers. *J Excip Food Chem* 2014;**5**:32–45.
58. Kolter K, Karl M, Gryczke A, Ludwigshafen am Rhein B. *Hot-melt extrusion with BASF pharma polymers: extrusion compendium*. Limburgerhof: BASF; 2012.
59. Huang S, O'Donnell KP, Keen JM, Rickard MA, McGinity JW, Williams III RO. A new extrudable form of hypromellose: AFFINISOL™ HPMC HME. *AAPS PharmSciTech* 2016;**17**(1):106–19.
60. Gosh I, Vippagunta R, Li S, Vippagunta SR. Key considerations for optimization of formulation and melt-extrusion process parameters for developing thermosensitive compound. *Pharm Dev Technol* 2012;**17**:9.
61. Liu H, Zhu L, Wang P, Zhang X, Gogos CG. Effects of screw configuration on indomethacin dissolution behavior in Eudragit E PO. *Adv Polymer Technol* 2012;**31**:12.
62. Fu J, Zhang L, Guan T, Tang X. Stable nimodipine tablets with high bioavailability containing NM-SD prepared by hot-melt extrusion. *Powder Technol* 2010;**204**:8.
63. Lyons JG, Blackie P, Higginbotham CL. The significance of variation in extrusion speeds and temperatures on a PEO/PCL blend based matrix for oral drug delivery. *Int J Pharm* 2008;**351**:8.
64. Crowley MM, Zhang F, Koleng JJ, McGinity JW. Stability of polyethylene oxide in matrix tablets prepared by hot-melt extrusion. *Biomaterials* 2002;**23**:4241–8.
65. Vervaet C, Remon JP. Continuous granulation in the pharmaceutical industry. *Chem Eng Sci* 2005;**60**:3949–57.
66. Gamlen M, Eardley C. Continuous extrusion using a raker perkins MP50 (multipurpose) extruder. *Drug Dev Ind Pharm* 1986;**12**:1701–13.
67. Dhenge RM, Cartwright JJ, Hounslow MJ, Salman AD. Twin screw granulation: steps in granule growth. *Int J Pharm* 2012;**438**:20–32.
68. Dhenge RM, Cartwright JJ, Hounslow MJ, Salman AD. Twin screw wet granulation: effects of properties of granulation liquid. *Powder Technol* 2012;**229**:126–36.
69. Iveson S, Litster J. Growth regime map for liquid-bound granules. American Institute of Chemical Engineers. *AIChE J* 1998;**44**:1510.
70. Vercruysse J, Toiviainen M, Fonteyne M, Helkimo N, Ketolainen J, Juuti M, et al. Visualization and understanding of the granulation liquid mixing and distribution during continuous twin screw granulation using NIR chemical imaging. *Eur J Pharm Biopharm* 2014;**86**:383–92.
71. Sayin R, El Hagrasy A, Litster J. Distributive mixing elements: towards improved granule attributes from a twin screw granulation process. *Chem Eng Sci* 2015;**125**:165–75.
72. Liu Y, Thompson M, O'Donnell K. Function of upstream and downstream conveying elements in wet granulation processes within a twin screw extruder. *Powder Technol* 2015;**284**:551–9.
73. Van Melkebeke B, Vervaet C, Remon JP. Validation of a continuous granulation process using a twin-screw extruder. *Int J Pharm* 2008;**356**:224–30.
74. Djuric D, Kleinebudde P. Impact of screw elements on continuous granulation with a twin-screw extruder. *J Pharm Sci* 2008;**97**:4934–42.
75. Vercruysse J, Burggraeve A, Fonteyne M, Cappuyns P, Delaet U, Van Assche I, et al. Impact of screw configuration on the particle size distribution of granules produced by twin screw granulation. *Int J Pharm* 2015;**479**:171–80.
76. Djuric D, Van Melkebeke B, Kleinebudde P, Remon JP, Vervaet C. Comparison of two twin-screw extruders for continuous granulation. *Eur J Pharm Biopharm* 2009;**71**:155–60.
77. Holm P, Jungersen O, Schaefer T, Kristensen H. Granulation in high speed mixers. I: effects of process variables during kneading. *Pharm Ind* 1983;**45**:806–11.
78. Tu W-D, Ingram A, Seville J. Regime map development for continuous twin screw granulation. *Chem Eng Sci* 2013;**87**:315–26.
79. Shah U. Use of a modified twin-screw extruder to develop a high-strength tablet dosage form. *Pharm Technol* 2005;**29**:52–66.
80. Vercruysse J, Díaz DC, Peeters E, Fonteyne M, Delaet U, Van Assche I, et al. Continuous twin screw granulation: influence of process variables on granule and tablet quality. *Eur J Pharm Biopharm* 2012;**82**:205–11.
81. Mu B, Thompson M. Examining the mechanics of granulation with a hot melt binder in a twin-screw extruder. *Chem Eng Sci* 2012;**81**:46–56.
82. Lakshman JP, Kowalski J, Vasanthavada M, Tong WQ, Joshi YM, Serajuddin A. Application of melt granulation technology to enhance tabletting properties of poorly compactible high-dose drugs. *J Pharm Sci* 2011;**100**:1553–65.
83. Keen JM, Foley CJ, Hughey JR, Bennett RC, Jannin V, Rosiaux Y, et al. Continuous twin screw melt granulation of glyceryl behenate: development of controlled release tramadol hydrochloride tablets for improved safety. *Int J Pharm* 2015;**487**:72–80.
84. Vasanthavada M, Wang Y, Haefele T, Lakshman JP, Mone M, Tong W, et al. Application of melt granulation technology using twin-screw extruder in development of high-dose modified-release tablet formulation. *J Pharm Sci* 2011;**100**:1923–34.
85. Kumar A, Gernaey KV, De Beer T, Nopens I. Model-based analysis of high shear wet granulation from batch to continuous processes in pharmaceutical production–a critical review. *Eur J Pharm Biopharm* 2013;**85**:814–32.
86. Barrasso D, Eppinger T, Pereira FE, Aglave R, Debus K, Bermingham SK, et al. A multi-scale, mechanistic model of a wet granulation process using a novel bi-directional PBM–DEM coupling algorithm. *Chem Eng Sci* 2015;**123**:500–13.
87. Bouffard J, Bertrand F, Chaouki J. A multiscale model for the simulation of granulation in rotor-based equipment. *Chem Eng Sci* 2012;**81**:106–17.
88. Eitzlmayr A, Khinast J. Co-rotating twin-screw extruders: detailed analysis of conveying elements based on smoothed particle hydrodynamics. Part 1: hydrodynamics. *Chem Eng Sci* 2015;**134**:861–79.
89. Barrasso D, El Hagrasy A, Litster JD, Ramachandran R. Multi-dimensional population balance model development and validation for a twin screw granulation process. *Powder Technol* 2015;**270**:612–21.
90. Nakatani M. Scale-up theory for twin-screw extruder, keeping the resin temperature unchanged. *Adv Polymer Technol* 1998;**17**:4.
91. Meyer R, Neu C, Smith-Goettler B. A decade's experience delivering clinical and commercial supplies using fully continuous hot melt extrusion. FDA-AIChE Workshop on Adopting Continuous Manufacturing, Methesda, MD; 2016.

92. Zhu Y, Chen B, Anderson BC. Extrusion simulation and its application to process scale-up. AIChE Annual Meeting, San Francisco, CA; 2013.
93. Dreiblatt A. Technological considerations related to scale-up of hot-melt extrusion process. In: Douroumis D, editor. *Hot-melt extrusion: pharmaceutical application*. New Delhi: John Wiley & Sons; 2012. p. 285–300.
94. Gao J, Walsh GC, Bigio D, Briber RM, Wetzel MD. Residence-time distribution model for twin-screw extruders. *AIChE J* 1999;**45**:2541–9.
95. Kumar A, Ganjyal GM, Jones DD, Hanna MA. Modeling residence time distribution in a twin-screw extruder as a series of ideal steady-state flow reactors. *J Food Eng* 2008;**84**:441–8.
96. Todd DB. Residence time distribution in twin-screw extruders. *Polymer Eng Sci* 1975;**15**:437–43.
97. Katharina P, Dirk L. Investigating process parameter mechanism for successful scale-up of a hot-melt extrusion process. https://tools.thermofisher.com/content/sfs/brochures/LR71-e-Investigating-Process-Parameter-Mechanism.pdf.
98. Kohlgrüber K. *Co-rotating twin-screw extruder*. Munich: Carl Hanser Verlag GmbH Co KG; 2012.
99. Markarian J. Compounders look to simulation software for savings in time and costs. *Plast Additiv Compoun* 2005;**7**:34–7.
100. Cunningham T, David C, Dannemiller D. Twin screw extruder simulation programs—what can they offer? *Plast Additiv Compoun* 2002;**4**:22–6.
101. Dubey SP, Abhyankar HA, Marchante V, Brighton JL, Blackburn K, Temple C, et al. Modelling and validation of synthesis of poly lactic acid using an alternative energy source through a continuous reactive extrusion process. *Polymers* 2016;**8**:164.
102. Zecevic DE, Wagner KG. Rational development of solid dispersions via hot-melt extrusion using screening, material characterization, and numeric simulation tools. *J Pharm Sci* 2013;**102**:2297–310.
103. Saerens L, Dierickx L, Lenain B, Vervaet C, Remon JP, De Beer T. Raman spectroscopy for the in-line polymer–drug quantification and solid state characterization during a pharmaceutical hot-melt extrusion process. *Eur J Pharm Biopharm* 2011;**77**:158–63.
104. MacPhail N, Meyer RF, Phillips JX, Gendron CM, Smith-Goettler B. NIR monitoring of a hot-melt extrusion process NIR monitoring of a hot-melt extrusion process. *Spectroscopy* 2011;**26**.
105. Almeida A, Saerens L, De Beer T, Remon JP, Vervaet C. Upscaling and in-line process monitoring via spectroscopic techniques of ethylene vinyl acetate hot-melt extruded formulations. *Int J Pharm* 2012;**439**:223–9.
106. Maniruzzaman M, Boateng JS, Snowden MJ, Douroumis D. A review of hot-melt extrusion: process technology to pharmaceutical products. *ISRN Pharm* 2012;**2012**:436763.
107. Tumuluri VS, Kemper MS, Lewis IR, Prodduturi S, Majumdar S, Avery BA, et al. Off-line and on-line measurements of drug-loaded hot-melt extruded films using Raman spectroscopy. *Int J Pharm* 2008;**357**:77–84.
108 Islam MT, Maniruzzaman M, Halsey SA, Chowdhry BZ, Douroumis D. Development of sustained-release formulations processed by hot-melt extrusion by using a quality-bydesign approach. *Drug Delivery and Translational Research* 2014;**4**:377–87.
109. Alig I, Fischer D, Lellinger D, Steinhoff B. 2005. Combination of NIR, raman, ultrasonic and dielectric spectroscopy for in-line monitoring of the extrusion process, Macromolecular symposia. Wiley Online Library, pp. 51–58.
110. Coates P, Barnes S, Sibley M, Brown E, Edwards HG, Scowen I. In-process vibrational spectroscopy and ultrasound measurements in polymer melt extrusion. *Polymer* 2003;**44**:5937–49.
111. Wang Y, Steinhoff B, Brinkmann C, Alig I. In-line monitoring of the thermal degradation of poly (l-lactic acid) during melt extrusion by UV–vis spectroscopy. *Polymer* 2008;**49**:1257–65.
112. Fang LY, Harris D, Krishna G, Moton AE, Prestipino RC, Steinman M, et al. High density compositions containing posaconazole and formulations comprising the same, Office, C.P., CA2720849 A1, 2009.
113. Desai D, Wong B, Huang Y, Ye Q, Tang D, Guo H, et al. Surfactant-mediated dissolution of metformin hydrochloride tablets: wetting effects versus ion pairs diffusivity. *J Pharm Sci* 2014;**103**:7.
114. Zu H, Yu P, Zhu L, Qiu Y. Feasibility study of manufacturing extended release hydrophilic matrix tablets via extrusion process. In: Poster, AAPS Annual Meeting, Orlando, FL, 2015.
115. Notario GF, Hom RC, Zhang J, Devcich KJ, Semla SJ. 2003. Extended release formulations of erythromycin derivatives. US Patent US 6551616 B1.
116. Qiu Y, Cheskin H, Engh K, Poska R. Once-a-day controlled-release dosage form of divalproex sodium I: formulation design and in vitro/in vivo investigations. *J Pharm Sci* 2003;**92**:8.
117. Bartholomaeus JH, Arkenau-Marić E, Galia E. Opioid extended-release tablets with improved tamper-resistant properties. *Exp Opn Drug Delivery* 2012;**9**:879–91.
118. Coperion ZSK101/Process Technology – Open Course.
119. <http://jmlordinternational.com/images/Extricom_screw_elements.pdf>.
120. Kohlgrber K. *Co-rotating twin-screw extruders, fundamentals, technology, and applications*. Cincinnati: Hanser Gardner Publications; 2008.
121. Sperling LH. *Introduction to physical polymer science*. 3rd ed Hoboken: John Wiley & Sons, Inc; 2001.
122. Ghebre-Sellassie I, Mollan MJ, Pathak N, Lodaya M, Fessehaie M, 2002. *Continuous production of pharmaceutical granulation*. US Patent 6499984, December 31, 2002.
123. Thompson MR, Weatherley S, Pukadyil RN, Sheskey PJ. Foam granulation: new developments in pharmaceutical solid oral dosage forms using twin screw extrusion machinery. *Drug Development and Industrial Pharmacy* 2012;**38**(7):771–84.
124. Li H, Thompson MR, O'Donnel KP. Progression of Wet Granulation in a Twin Screw Extruder Comparing Two Binder Delivery Methods. *AICHE Journal* 2015;**61**(3):780–91.
125. Muehelenfeld C, Thommes M. Miniaturization in pharmaceutical extrusion technology: feeding as a challenge of downscaling. *AAPS PharmSciTech* 2012;**13**(1):94–100.
126. Schæfer T. Growth mechanisms in melt agglomeration in high shear mixers. *Powder Technology* 2001;**117**(1-2):68–82.
127. Juul Thomsena L, Schaefera T, Kristensena HG. Pellets Prepared by Melt Pelletization II. Hydrophobic Substances as Meltable Binders. *Drug Development and Industrial Pharmacy* 1994;**20**(7):1179–97.
128. Voinovich D, Moneghini M, Perissutti B, Franceschinis E. Melt pelletization in high shear mixer using a hydrophobic melt binder: influence of some apparatus and process variables. *European Journal of Pharmaceutics and Biopharmaceutics* 2001;**52**(3):305–13.
129. Keleb EI, Vermeire A, Vervaet C, Remon JP. Twin Screw Granulation as a Simple and Efficient Tool for Continuous Wet Granulation. *International Journal of Pharmaceutics* 2004;**273**(1-2):183–94.
130. Cartwright JJ, Robertson J, D'Haene D, Burke MD, Hennenkamp JR. Twin Screw Wet Granulation: Loss in Weight Feeding of a Poorly Flowing Active. *Pharmaceutical Ingredient Powder Technology* 2011;**238**:116–21.
131. <http://sawyerhanson.com/uploads/docs/FlexWall%20Classic%20Feeder%20Brochure.pdf>.
132. Carley JF, Mckelvey JM. Extruder scale-up theory and experiments. *Ind Eng Chem* 1953;**45**:989–92.
133. Brown C, et al. HME for Solid Dispersion: Scale-up and Late-Stage Development. In: Shad N, editor. *Amorphous Solid Dispersion: Theory and Practice*. Springer; 2014.
134. Schæfer T, Mathiese, C. Melt pelletization in a high shear mixer. VII. Effects of product temperature. *International Journal of Pharmaceutics* 1996;**134**(1-2):105–117

CHAPTER

32

Development, Scale-Up, and Optimization of Process Parameters: Roller Compaction Theory and Practice

J.M. Rowe[1], *S.T. Charlton*[2] *and R.J. McCann*[3]

[1]Bristol-Myers Squibb, New Brunswick, NJ, United States [2]Bristol-Myers Squibb, Moreton, Wirral, United Kingdom
[3]AstraZeneca, Mt. Vernon, IN, United States

32.1 INTRODUCTION

Roller compaction is the preferred method for dry granulation in the pharmaceutical industry. It is a continuous manufacturing technique introduced in the second half of the 19th century in the coal industry to compact powdered coal into briquettes.[1] The technology was later adopted by metallurgists for roll pressing of metal powders before finally being introduced into the food, agricultural, and pharmaceutical industries.[2] The pharmaceutical industry has traditionally been very good at adopting technologies from other industries and learning from the prior knowledge. However, the behavior of metals and organic materials differ dramatically, so significant efforts have been made, and are still ongoing, to adapt the process for pharmaceutical applications.

Good powder flow is particularly important when manufacturing tablets, the most common solid oral dosage form, because rapid and accurate filling of the tablet die is essential to achieve consistent tablet weights, and thus, consistent drug dose. When active pharmaceutical ingredients (API) possess desirable flow characteristics and demonstrate minimal tendencies toward segregation, direct compression processes, which involve the compaction of ungranulated powders into tablets, are feasible. However, API powders typically possess poor flow characteristics, which present handling challenges during the manufacturing process and are a potential source for segregation in powder blends. Pharmaceutical manufacturing processes utilize roller compaction operations to mitigate these challenges caused by poorly behaving API by improving flow and content uniformity of powder blends. The roller compaction technique densifies powder blends to produce granules with particle sizes larger than that of the input material, and thus improved flow characteristics compared with ungranulated material. In addition, the granulation process reduces the potential for API segregation, which can result in poor API content uniformity. Fig. 32.1 demonstrates where the roller compaction operation would fit into a typical manufacturing process for tablet dosage forms. Since tablets are the most common oral solid dosage form, the remainder of the chapter assumes tablets as the final dosage form, unless otherwise noted.

Compared with direct compression processes, roller compaction operations are better suited to handle powders with the following attributes: (1) small particle sizes and/or challenging morphologies (eg, high aspect ratio particles), (2) low bulk density, (3) poor flow properties, and (4) high risk for API segregation. Consequently, roller compaction workflows are more capable of accommodating variability in API properties compared with direct compression processes. Roller compaction is a dry process; therefore it is more suitable for moisture- and heat-sensitive material, which is an advantage over wet granulation. Furthermore, roller compaction is a continuous process, enabling a single instrument configuration able to accommodate a greater range of batch sizes compared with wet granulation.

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00032-7

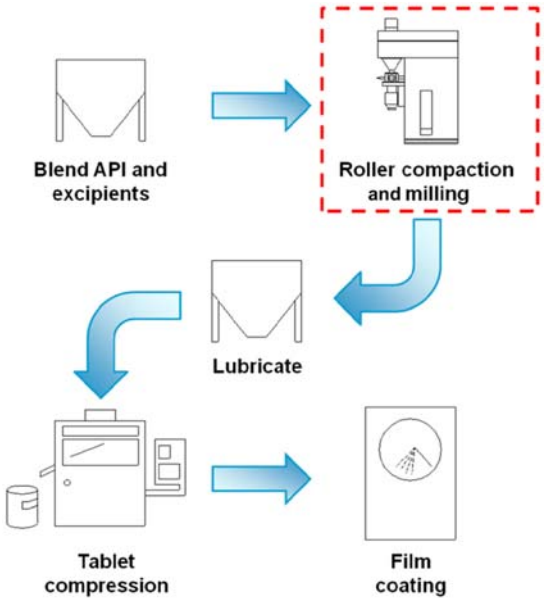


FIGURE 32.1 Schematic showing where roller compaction fits into the manufacture process for tablet dosage forms.

Benefits of the roller compaction process include

- *Improved flow*—larger granules typically flow better than fine powders, resulting in improved powder transfer, flow in hopper, and die-filling during tabletting
- *Increased bulk density*—faster bulk transfer; reduced dust generation; smaller bins can be used during post-granulation blending; reduced in-die fill volume during tabletting, facilitating manufacture of large tablets
- *Improved blend uniformity*—contents of blend "trapped" into granules, facilitating uniform distribution of materials throughout the entire blend and reducing potential for materials to segregate; this is particularly pertinent to API distribution and subsequent dose uniformity
- *Continuous process*—enables manufacture of larger batch sizes by simply increasing operating times, which allows a single roller compaction unit to support a wide range of batch sizes (eg, a typical pilot scale roller compactor can manufacture batch sizes of 1–200 kg)
- *Improved dust control*—consolidation and densification of powder into granules reduces the amount of fines and discrete API particles, reducing dust generation during tablet manufacture and exposure risk to operators
- *Improved tablet disintegration and dispersal*—granules readily absorb water, enabling faster disintegration compared with tablets made by direct compression; immediate release tablets: disintegrate to granules, which disperse and disintegrate themselves; capsules: granules improve dispersion and mitigate plug formation

This chapter will provide an overview of roller compaction and discuss key elements of the operation, including equipment design, simulation, and characterization tools that facilitate efficient execution of roller compaction operations, trouble-shooting methods, scale-up considerations, as well as review current modeling techniques that simulate the roller compaction process and improve mechanistic understanding of the process. Case studies will be used throughout the chapter to highlight discussions in the text.

32.1.1 Overview of roller compaction materials and operation

The overarching goal of the roller compaction process is to transform poorly flowing, fine powders into granules of a desired particle size range and distribution with appropriate flow and compactability. Optimal particle size distribution also facilitates good granule flow, which is important for efficient transfer of material into hoppers and powder feed systems to ensure accurate and reproducible metering of doses (ie, capsule or sachet filling, and filling of tablet dies on a high-speed rotary press). Achieving optimal particle size distribution is important with respect to tablet weight uniformity, especially for low-weight tablets that may be susceptible to significant weight variation if granules are too large. Granule compactability is essential for the manufacture of tablets, which must be of sufficient hardness and tensile strength to prevent damage during film coating, packaging and shipping.

A basic overview of the roller compaction process is depicted in Fig. 32.2. Briefly, loose powder is fed into the roller compactor from a feed system, which typically sits atop the instrument. The feed system is often comprised of an agitator within the hopper and one or more feed screws that convey the powder from the hopper to the rolls, and also facilitates particle and air removal from the powder bed. An integrated vacuum deaeration system, consisting of a sintered metal filter or mesh screen, may be incorporated into the feed system to better remove air from the powder bed (deaeration). This deaeration helps to reduce the friability of the compacted ribbon edges and leads to denser ribbons.[3,4] The powder is passed from the feed system between the two counter-rotating rolls of the roller compactor. The rolls apply a compressive force to the powder, compressing the material to form a densified compact, typically in the form of a ribbon (sheet). These compacts then typically undergo a size reduction process (milling) to produce free-flowing granules that are amenable for

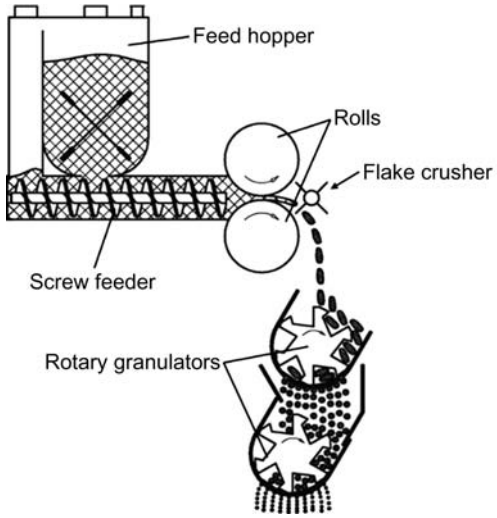


FIGURE 32.2 Schematic illustrating a typical roller compaction process. *Source: Courtesy from Alexanderwerk AG.*

downstream processing, such as capsules or tablets.[5] Ultimately, the success of a roller compaction process is assessed on the properties of the milled granules and the final dosage form, such as tablets.[6,7]

It should be noted, however, that designs differ between manufacturers and models, which will be addressed in more detail later in the chapter. In Fig. 32.2, the feed screw is oriented horizontally with vertically aligned rolls; however, alternate configurations include a horizontal or angled roll orientation, depending on the manufacturer and instrument model. Furthermore, a number of different roll surfaces are commercially available, including smooth, knurled, ribbed, and serrated. For larger scale roller compactors that include an integrated mill system, the milling apparatus is located immediately after the rolls, and may include a flake crusher followed by a one- or two-stage mill, which is typically a screen and an impeller operating in a rotary or oscillating motion. Smaller lab-scale instruments do not usually include mills, and thus the ribbons must be milled using a separate unit.

In today's quality by design (QbD) paradigm, it is important to establish a robust design space of process parameters and material properties. QbD strategies facilitate a more comprehensive understanding of how material and process attributes impact product quality and performance, as well as providing insight into potential interactions between the two.[8]

32.1.1.1 Material considerations

A typical roller compaction formulation consists of API, filler (eg, microcrystalline cellulose (MCC) and/or lactose), disintegrant (eg, croscarmellose sodium, crospovidone), lubricant (eg, magnesium stearate, stearic acid), and often times a glidant (eg, silicon dioxide) to reduce potential friction with the rolls and facilitate flow of powder. Formulations possessing drug loadings <60% wt are generally found to be amenable for roller compaction. Higher drug loading formulations generally have a higher tendency for material adhesion to the rolls, which adversely impacts ribbon quality and material throughput, thus presenting challenges to maintaining a consistent flow of powder to the rolls. One way to mitigate material-related challenges typically associated with API powder properties is through the appropriate selection of excipients. Specific grades of excipients, such as lactose and MCC, are available that possess optimized physical properties in terms of flow and compactability, and thus can be used to improve the overall powder properties of an active blend. A proposed list of desirable properties for blends intended for roller compaction is detailed in Table 32.1.[9]

32.1.1.2 Key operating parameters and product attributes

A list of key operating parameters and product attributes for the roller compaction process are given in Table 32.2.

The key parameters typically monitored during roller compaction operations are roll gap, feed screw speed, and roll pressure because they most impact ribbon/granulation quality. In most pilot-scale and larger roller compactors, a floating roll is utilized to minimize fluctuations in ribbon solid fraction caused by variations in powder flow, and may include a closed feedback loop to modify the feed screw speed in response to inconsistencies in powder flow. Therefore, roll gap and feed screw speed often fluctuate continuously throughout the roller compaction operation due to variations in powder flow and a feedback loop controlling feed screw speed, although the average of these values should remain consistent once the operation has reached steady state. For example, when roll pressure and roll-gap targets are defined, then feed screw speed is varied in response to changes in roll gap (due to variation in powder flow) to ensure that sufficient powder is delivered to the rolls to maintain the target gap. Roller compactors with two fixed rolls, as opposed to a floating roll design, tend to show greater variability in granule properties as a result of powder flow variation. Additional process parameters, such as roll speed and mill screen size may also impact ribbon/granule properties, but are generally considered to be "fixed" parameters for a given run because they do not fluctuate during the roller compaction operation. Roll speed selection is often based on throughput considerations. Changes in roll speed may change ribbon tensile strength for a given roll pressure, because the dwell time (ie, the duration that pressure is applied to the powder by the rolls) is impacted. Mill screen size will influence granule particle size, flow, and

TABLE 32.1 Properties of an Ideal Roller Compaction Material

Property	Value	Target value
Compactability	At 0.7 solid fraction	>1 MPa
Loss of compactability[a]	At <0.9 solid fraction	>1.7 MPa
Bulk density	At >0.14 solid fraction	>0.2 g/mL[b]
Wall friction	Angle of wall friction	<20 degrees[c]
Flow assessment	Flow function coefficient	>4[c]
	Carr's Index	<35%
Solid-state properties	Melting point	>90°C
	Glass transition (T_g)	>90°C
	Loss of crystallinity during compaction	None
Particle size for content uniformity	Blend potency	Meets Rohr's criteria[10]
Stability with excipients	Acceptable stability	>2 years at room temp[11]
Material adhesion	Not yet determined	Not yet determined

[a]*Loss of compactability defined as the compactability remaining in granules.*
[b]*USP <616> Method I—Measurement in a graduated cylinder.*
[c]*Shear cell method, measured over the range 1 to 5 KPa.*
Adapted from Leane et al. A proposal for a drug product Manufacturing Classification System (MCS) for oral solid dosage forms. Pharm Dev Technol *2015;**20**(1):12–21.*

TABLE 32.2 Key Product and Operating Parameters Typically Associated With Roller Compaction

Key operating parameters	Ribbon attributes	Granule attributes
Roll pressure	Solid fraction/ porosity	Particle size distribution
Feed pressure (ie, feed screw speed)	Tensile strength	Compactability
Roll speed	Thickness	Flow
Roll gap		

compactability, and a change of mill screen size may necessitate changes to other process parameters. Thus, appropriate selection of these "fixed" parameters should be conducted early in the development process.

Key product attributes typically monitored during the roller compaction operation focus on ribbon properties because they provide a good indication of the resultant granule properties. Generally, from batch to batch, ribbons possessing similar solid fraction, tensile strength, and thickness will yield granules of comparable particle size distribution, flow, and compactability properties.

32.1.1.2.1 Ribbon solid fraction

Ribbon solid fraction is commonly used as an in-process target because it provides a strong indication of granule properties. Solid fraction, which describes the extent of compaction of a material, is defined as the ratio of a material's envelope density to its true density (ρ/ρ_{true}). Porosity is the inverse of solid fraction; therefore a solid fraction of 1 indicates that the compact has zero porosity. Solid fraction is proportional to the force applied to the powder, and is independent of ribbon thickness. There is a positive correlation between solid fraction and tensile strength, with ribbons of greater solid fraction having a greater tensile strength. If the solid fraction (or tensile strength) of ribbon is too low, then the ribbon will break up to fine powder rather than granules when milled, resulting in a poorly flowing product. Ribbons of adequate solid fraction typically produce larger, better-flowing granules; however, overly compacted granules exhibit poor compactability during tablet manufacture. Therefore, the correct balance between fines content generated during roller compaction is necessary to achieve granules with both proper flow and compactability characteristics (Fig. 32.3). The optimal ribbon solid fraction range will be product specific to a certain extent, but the ideal solid fraction for ribbon is generally in the range of 60–80%. Fig. 32.4 illustrates typical solid fractions ranges for ribbons, with respect to that for loose powders and tablets.

Granule compactability is not a major concern when manufacturing granules for filling into capsules or sachets, therefore it may be more desirable to produce granule of a higher solid fraction, in the range 75–85%, to achieve improved granule flow.

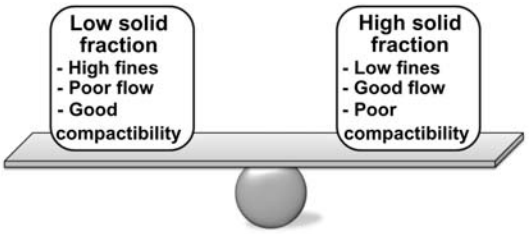


FIGURE 32.3 Getting the balance right—ribbon solid fraction impacts granule particle size distribution, flow, and compactability.

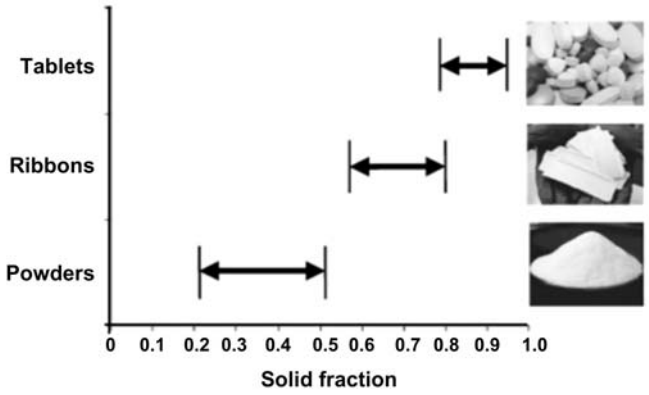


FIGURE 32.4 Typical solid fractions of pharmaceutical materials.[26] *Source: ©2004 with permission from Elsevier.*

In this instance, the target ribbon solid fraction range should also be assessed in relation to dissolution performance.

32.1.1.2.2 Ribbon tensile strength

Ribbon tensile strength is derived from the breaking force and thickness of the ribbon, and has a positive correlation with solid fraction. Greater ribbon tensile strengths can be achieved by increasing roll pressure and, to a lesser extent, slowing the roll speed (ie, increase the dwell time). Ribbon tensile strength may be characterized using a three-point bend test (details discussed in Section 32.2) to measure the ribbon breaking force, which can then be translated into a tensile strength. Low tensile-strength ribbons more readily reduce to fine powder during milling, which may have a negative effect on subsequent blend flow and negates the advantages borne from granulation. However, over-compressed ribbons typically produce tablets of lower tensile strength, which is attributed to reduced inherent compaction potential after granulation.[12] Recommended tensile strengths for roller compaction ribbons are >1 MPa at 0.7 solid fraction[9] to ensure that robust granules are produced. Ribbon of higher tensile strength may be preferred if the granule will not be subsequently recompacted, as in the case of granules intended for capsules.

32.1.1.2.3 Roll gap and ribbon thickness

Roll gap at the point of maximum compaction essentially dictates ribbon thickness (although material relaxation after the ribbon is released from the rolls will cause a slight increase in ribbon thickness compared with roll gap values). Fluctuations in powder flow may affect in-process variation of roll gap and ribbon thickness due to changes in the mass of powder passing between the rolls. Roller compactors that utilize a floating roll will respond to changes in powder flow by dynamically altering roll gap to maintain a constant solid fraction. Changes in roll gap should not impact ribbon tensile strength if powder density is uniform throughout the sample, because tensile strength should increase proportionally with increasing ribbon thickness. However, there is evidence to suggest that density is not uniform through the thickness of the ribbon, especially for thicker ribbons, with material being more compacted at the upper and lower surfaces and less compacted in the center. Thus, thicker ribbons may present a higher tensile strength, but result in a higher proportion of fines.[13–17] Therefore, ribbon thickness may influence granule particle size distribution, flow, and compactability, especially when using a larger roll gap.

32.1.1.3 Operating principles

Plastically deforming materials undergo a finite amount of compaction and deformation, and there is a limit to the degree of fracturing a brittle material will incur under typical pressures applied during roller compaction and tablet manufacture.[18] To relate the limits of a material's compaction and deformation to measurable tablet properties, the effect of recompression (subjecting a material to multiple compression steps, ie, roller compaction followed by tablet manufacture) on the compactability of a material has been well documented. Recompression generally causes a reduction in tablet strength, and this reduction is more significant when the initial compaction is carried out at a higher pressure,[19–25] as demonstrated in Fig. 32.5. This loss in tablet strength is often referred to as a "loss of compactability" or "reduction in strength" and has been attributed to: (1) work hardening of granules and (2) granule size enlargement.[24,25] The concept of "work hardening" was proposed as an empirical method to describe the magnitude of the reduction in compactability; it can be defined as the irreversible plastic deformation of particles that introduces defects that "harden particles."[23] For the latter cause, larger particles (formed from the first compaction step) are hypothesized to reduce compact strengths because they possess lower surface areas, thus reducing packing efficiency and availability for bonding.[24]

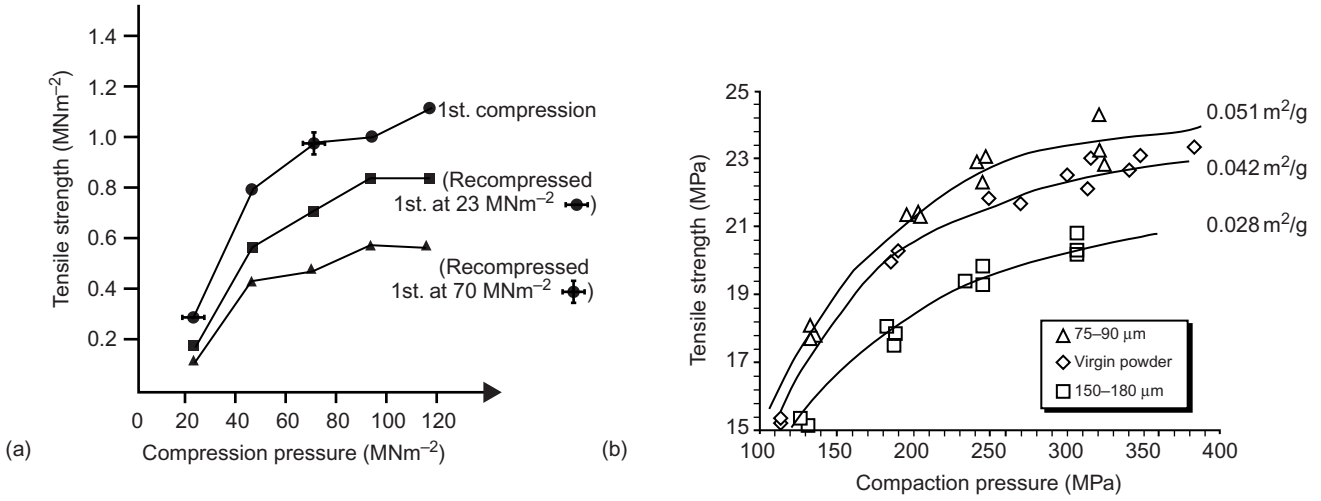


FIGURE 32.5 Effect of recompression on the tensile strength on (a) Starch 1500[23] and (b) MCC Avicel PH-102.[24] *Source: (a) ©1983 with permission from Taylor and Francis and (b) ©2006 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.*

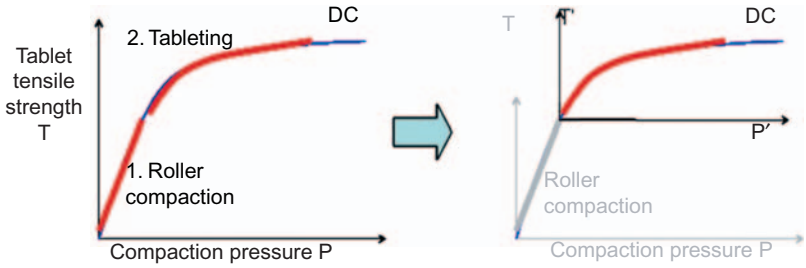


FIGURE 32.6 Proposed relationship between direct compression (*DC*) and roller compaction (*RC*) curves.[19] *Source: ©2008 with permission from Elsevier.*

Farber et al. presented a relationship correlating roller compaction conditions and tablet strength, in which compaction is assumed to be cumulative throughout the lifetime of a material.[19] This theory, coined as the "unified compaction curve model," postulates that compact strength (ribbon and tablet) is generated irreversibly, as if strength is primarily controlled by plastic deformation of primary particles. Therefore, the tensile strength of compacts manufactured from granules is dependent on the inherent compactability remaining after granulation. As such, if a relatively large amount of particle bonding and tensile strength is generated during roller compaction, then there is less potential for gaining tensile strength during tablet compaction, resulting in weaker tablets. When the total compaction history is accounted for, the compaction behavior of uncompacted blends and roller-compacted granules ultimately adheres to a single master compaction curve, or unified compaction curve. The concept of the unified compaction curve is illustrated in Fig. 32.6. Application of the unified compaction curve model to an MCC/lactose system is demonstrated in Fig. 32.7. It should be noted that the model is generally more applicable to systems containing a significant proportion of plastically deforming material, such as MCC and starch.

32.1.1.4 Roller compaction equipment design

There are many different roller compactor designs in the market today, and examples of the various designs are given next (Fig. 32.8). The main differences between the roller compactor designs are the roll orientations, feed screw systems, roll surfaces, and mill types.[14,27,28] The rolls may be configured in parallel, perpendicular, or at 45 degree to the horizontal, in addition to related differences in the feed system orientation. Feed systems may vary by the feed screw orientation and number of feed screws (Fig. 32.9). Possible orientations for feed systems, based on current commercially available units

include: (1) vertical gravity or vertical feed screw (Freund-Vector (Marion, IA), (2) horizontal feed screws (Alexanderwerk (Remscheid, Germany)), (3) a horizontal conveying screw followed by a vertical tamping screw (Fitzpatrick (Elmhurst, IL)), and (4) horizontal feed screw followed by an angled tamping screw (Gerteis (Jona, Switzerland)). The specific arrangement can influence compaction behavior within the roller compactor, such as feed pressures. Roller compactors with a horizontal feed screw and vertically stacked rolls typically come equipped with either a single or double feed screw to convey the powder to the rolls. The use of single- or double-feed screws is dependent upon equipment scale and roll width. For example, the pilot plant scale Alexanderwerk WP120 roller compactor uses a single-feed screw, due to the narrow roll width of 40 mm, versus the use of a double-feed screw on the production scale WP200, which uses 75 mm wide rolls. Consistent delivery of powder to the entire roll width is key; so for larger roll widths, the use of a single feed screw is not common because the periodicity of a single-feed screw can be observed in the widthwise density distribution of the roller compacted ribbon.[29] The use of a double-feed screw can help reduce this widthwise density variability by moving the maximum density achieved for each feed screw closer to the edge of the roll.[30]

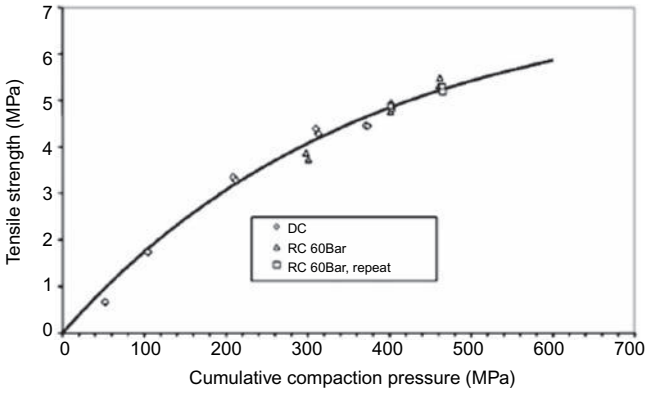


FIGURE 32.7 Unified compaction curve for a 1:1 MCC:Lactose blend (solid line). Direct compression (*DC*) compaction and roller compaction (*RC*) experimental data points denoted for indicated conditions.[19] *Source: ©2008 with permission from Elsevier.*

Deaeration of incoming uncompacted feed powder is a common component of the feed screw system. The powder being fed to the rolls has an apparent density equivalent to the powder's bulk density. Therefore, air is entrapped within the powder mass being conveyed to the rolls. Entrapped air increases the propensity of roller compacted ribbons to laminate once the compressive force from the rolls is removed. Removal of entrapped air within the feed powder improves the powder's compactability. The air can be removed via gravity, a tamping screw, or a

FIGURE 32.8 Commercially available roller compactors. *Source: Courtesy: Freund-Vector Corporation, The Fitzpatrick Company, Gerteis Maschinen + Process engineering AG, Alexanderwerk AG.*

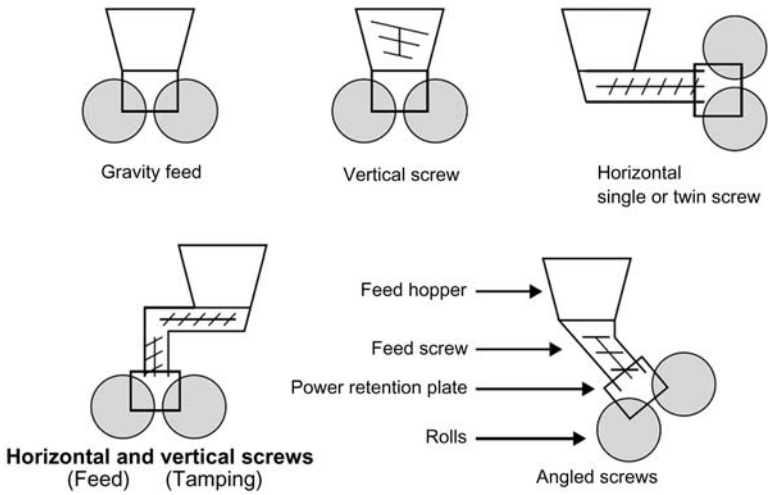


FIGURE 32.9 Feed system orientations.[12] *Source: ©2012 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.*

vacuum depending on the design of the roller compactor. Vacuum deaeration is most effective way and is accomplished through the use of sintered metal filters and a vacuum pump. The filter is typically incorporated into the walls of the feed screw system and allows the powder to flow across the inner surface of the filter. The vacuum pump is then connected to the outer surface and pulls air out of the powder. This increases the density of the powder and prepares the material for compaction at the rolls.

A variety of roll surfaces are available for the various roller compactors; the surfaces can vary from polished smooth (flat) rolls to textured rolls with knurling, as shown in Fig. 32.10.[31] Materials with a tendency to stick to surfaces may benefit from testing with different types of roll surfaces. In addition to assessing the impact of roll surface, these cases may require further exploration of lubrication levels and types of lubricant (ie, magnesium stearate vs stearic acid, sodium stearyl fumarate), as well as benefit from the use of external lubrication units to mitigate sticking challenges.[32]

Some rolls are interlocking punch-and-die while others are pocketed to create briquettes.[33] The different roll surfaces are used to alter the amount of friction between the powder particles and roll to facilitate the powder being pulled through the rolls to create a ribbon. A punch-and-die roll arrangement consists of one roll that fits inside the other, which removes the need for side seals to contain the powder during roller compaction. This arrangement is commonly associated with the use of Gerteis roller compactors, but the concept can be adapted to other equipment designs as this roll configuration yields significantly different ribbon density distributions compared with more traditional roll configurations, producing ribbon in which the density is highest at the edges of the ribbon and lowest in the center. In more traditional roll configurations that utilize a static side seal, frictional drag effects between the powder and side seal minimize powder accumulation at the edges of the roller compacted ribbon, yielding lower density edges with a higher density center. Although the density distribution within the ribbon may be different, the same average ribbon density can be achieved for both arrangements by adjusting compression force applied by the rolls. It is important to note, however, that the resultant milled particle size distribution is a function of the density of the ribbon coming off the rolls; thus, there may be a shift in the amount of fines and coarse particles depending on the proportion of low- and high-density regions within the ribbon.

The manner in which ribbons are milled is important because ribbons are not the final dosage form.[34] The mill design is the final differentiation among roller compactor equipment manufacturers. Some manufacturers use an integrated mill with the roller compactor, which consists of a single (Gerteis (Jona, Switzerland)) or double (Alexanderwerk (Remscheid, Germany)) set of screens with rotating impellers. The other option is an external mill (Fitzpatrick (Elmhurst, IL)) if the roller compactor does not have an integrated mill. The choice of mill screen size, rotor speed, and rotor-to-screen clearance will impact the final granule particle size distribution. Optimization of the milling operation to ensure proper granule flow and particle size distribution for blend uniformity should be performed as part of process development.

Roller compactor equipment design can have significant impacts on the ribbon products. For example, roller compactor design elements, such as roll design and feed system, have been shown to cause roller compacted ribbons to exhibit a three-dimensional density distribution. The density distribution of roller compacted ribbons has been characterized in previous

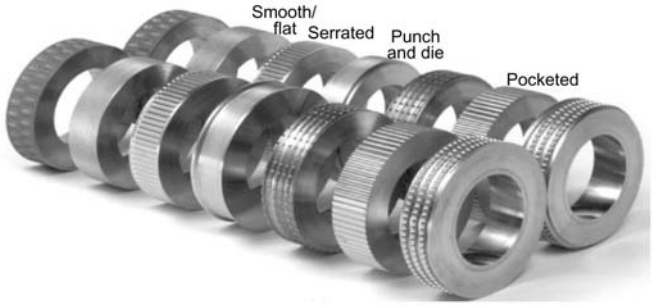


FIGURE 32.10 Roll design options.[12] *Source: Copyright © 2012 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.*

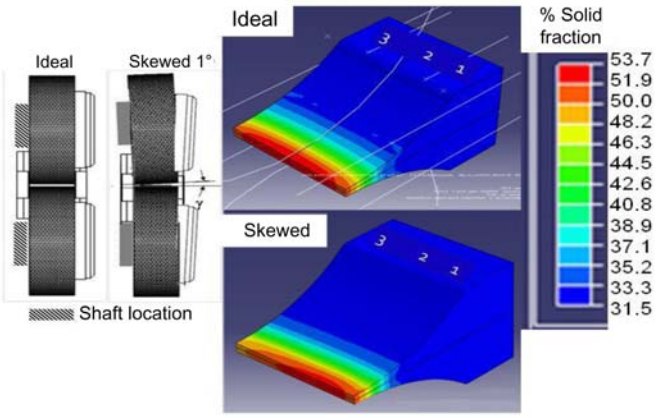


FIGURE 32.11 Finite element modeling explains observed anisotropic density distribution in ribbons.[38]

studies both experimentally and computationally for metals and organic materials.[14,30,35–37] These studies have shown that ribbons coming off of a roller compactor generally exhibit a parabolic widthwise density distribution, where the edges possessed a lower density than the center of the ribbon for rolls equipped with fixed side walls.[30,35] The lower density edges of the roller compacted ribbons were attributed to frictional effects between the powder and fixed side walls.[35] The lengthwise density variation was shown to manifest itself as a sinusoidal wave of higher density running throughout the ribbon's length.[14,37] Therefore, ribbons with the same average density do not necessarily perform the same in downstream processes. In support of these studies, McCann[38] used near infrared (NIR) spectroscopy to map the density distribution of ribbons and found that the ribbon edge closer to the machine body had a slightly higher density than the outer ribbon edge, in addition to the observation that the maximum ribbon density was offset from the center line. Finite element computational modeling was used to show how the cantilevered design of the roller compactor used in the study contributed to the anisometric ribbon density. The upper roll drive shaft was found to have deflected from parallel, which can cause the outer edge of the roll to deviate at a small angle away from being parallel with the bottom roll. In the computational model, a 1 degree deviation from parallel was simulated. This small deviation from parallel showed that the applied pressure was anisotropic with the inner ribbon edge, experiencing higher pressure than the outer edge, as shown by the skewed orientation in Fig. 32.11.

32.2 IN-PROCESS ANALYTICAL CHARACTERIZATION TOOLS

Several tools, listed in Table 32.3, have been used to improve process understanding of the roller compaction operation, as well as to facilitate accurate characterization of product qualities. In the table, bench-top roller compactors are mentioned because they have the potential to serve as a material-sparing option during the design and optimization of roller compaction operations. Commercially available bench-top roller compactors include the TFC-LAB micro-roll compactor (Freund-Vector), Fitzpatrick IR 220 Chilsonator (Fitzpatrick), BT120 (Alexanderwerk), and the Miniroll roller compactor (Riva (Hampshire, UK)). Input requirements for these bench-top units can be as low as 5 g of material. However, due to the absence of peer-reviewed literature at this time regarding the contribution of these units to the general mechanistic or scale-up knowledge for roller compaction, the bench-top units will not be discussed in this chapter.

TABLE 32.3 List of Tools and Tests Used to Characterize the Roller Compaction Operation

Tool/test	Application
PROCESS DEVELOPMENT/UNDERSTANDING	
Instrumented roll	Modeling
	Mechanistic understanding
Compaction simulator	Roller compaction simulation
Bench-top roller compactors	Roller compaction simulation
	Small-scale roller compaction production
CHARACTERIZATION	
Envelope density analyzer	Ribbon density
Three-point bending flexural test	Ribbon strength
Arc punch	Ribbon density
Laser profilometer	Ribbon density
Process analytical technology tools	Ribbon density, tensile strength, moisture content, API content, granule size

32.2.1 Instrumented roll

The current pharmaceutical industry standard roller compactors include operator interfaces that allow for online monitoring of operating conditions, including roll pressure. However, the roll pressure readings do not accurately reflect the actual compaction pressures experienced by the powder because they include equipment effects. To address this gap, the instrumented roll has been developed as a means to monitor the compaction pressure of powders in-situ during the roller compaction process.[35,39–43] This technology has been adapted from the metallurgy field, where measurements of normal forces, or pressures, on materials compacted during a rolling process have been reported as early as 1933.[44–46] Fig. 32.12 shows a roller compactor roll that has been instrumented with three sensors, positioned along the full width of the roll. Pressure sensors embedded on the surface of the roll of a roller compactor are used to measure the pressure exerted by the rolls onto powder during the roller compaction process. Thus, for this particular instrumented roll, pressure readings may be recorded along the width of the roll for every rotation, enabling measurements at the roll center and the sides, where edge effects have been reported. It should be noted that instrumented roll data reported in literature are not collected in a uniform manner because the instrumented roll is not a standardized tool and is often specially assembled in-house. Variability in instrumented roll specifications exists in terms of sensor placement, type (eg, strain gauge load cells[35,41–43,47] or piezoelectric transducers[17,39,40,48–50]), and number of sensors.

The ability of the instrumented roll to monitor normal and shear stresses in-situ has provided scientists with an in-depth look into the roller compaction operation, enabling a greater mechanistic understanding of the process and facilitating the development of predictive models through the establishment of empirical relationships and compaction theories. The basic function of this tool is to map the evolution of roll pressure and stress profiles during the roller compaction process, from which greater knowledge evolves through the quantification and prioritization of relationships between operating conditions and product quality.[35,50] Data collected from instrumented rolls are in the form of pressure and stress profiles, as shown in Fig. 32.13. From these plots, maximum compaction pressures (P_{max}) and shear stresses are extracted to provide insight into relationships between operating conditions and experimental parameters, including formulation, process, and equipment considerations, to compaction quality.[17,35,43,49,51]

32.2.1.1 *Mechanistic understanding*

One of the key drivers behind the use of instrumented roll technology was to increase the mechanistic understanding of the roller compaction operation by providing information on the actual pressure that imparts powder compression, allowing for deconvolution of ribbon quality from equipment and material influences. Instrumented roll data have substantiated strong correlations between ribbon solid fraction (γ_R) and P_{max} and have demonstrated the independence of this relationship from equipment factors (ie, roller compactor type/scale), processing conditions, or material selection.[19,26,52–54] In Reynolds et al.,[52] P_{max}- γ_R profiles for an active blend, where drug load exceeded 30%, lie virtually along the same curve despite varying

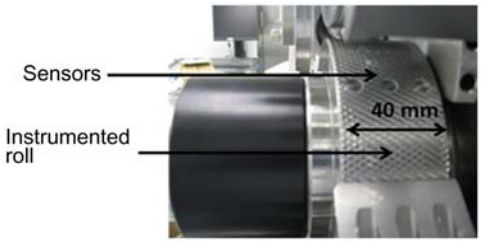


FIGURE 32.12 Roller compactor roll that has been instrumented with three pressure (strain gauge load cell) sensors.[42] *Source: ©2012, with permission from Elsevier.*

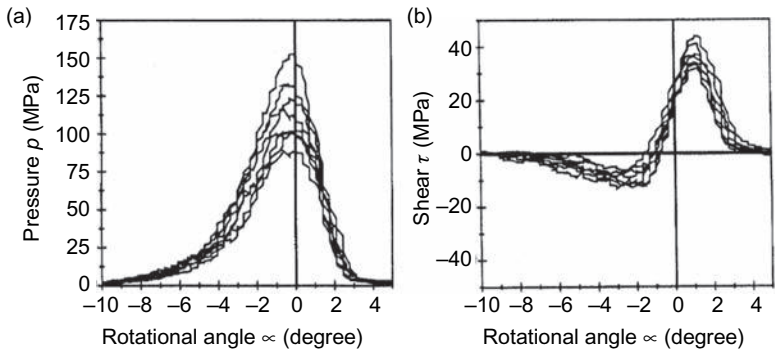


FIGURE 32.13 Typical data collected by instrumented roll: (a) Pressure profile as a function of rotational angle, α, for quartz (b) Shear stress profile as a function of α for quartz.[43] *Source: ©2002, with permission from Elsevier.*

operating parameters, operating scale, and even slight changes to the formulation (Fig. 32.14). Hydraulic roll pressure, screw speed, and roll speed were varied between 20 and 60 bar, 35 and 75 rpm, and 5 and 10 rpm, respectively, in this study. Intragranular lubricant levels were also adjusted between 0.25% and 0.75%, by weight.[52] However, regardless of the wide range of input conditions examined in this work, the resultant P_{max} ultimately dictated γ_R.

Given the importance of P_{max} to product quality, numerous studies have been conducted to elicit trends regarding the impact of different operating parameters, such as roll gap, roll speed, and feeding conditions (ie, feed screw speed) on compaction pressure (Fig. 32.15). This information is highly valuable because it enables a more thorough understanding of the design space for the roller compaction operation and imparts process robustness, providing confidence in product quality in the event that operating conditions drift during manufacture. It is important to note that that these trends are unique to each system and therefore must be re-established for each material/equipment combination. Guigon et al. used the instrumented roll to determine that a constant P_{max} was achieved during the roller compaction of hydragilitte (SH100), lactose, and sodium chloride powders through the maintenance of a constant roll gap (Fig. 32.15a).[14] Constant roll gap was achieved by controlling feed screw and roll speeds, although feed-screw speeds played a more dominant role in impacting roll gap, and thus P_{max}. Yu et al. demonstrated that P_{max} was relatively independent of roll speed when roller compacting MCC, but was highly dependent for a dibasic calcium phosphate dihydrate (DCPD) system (Fig. 32.15b).[50] Yusof et al. reported a sizeable dependency of P_{max} on feed rate when roller compacting maize powder (Fig. 32.15c).[55] However, roll gap and roll speed were found to have a larger influence on P_{max} than feed screw rate. Ingelbrecht et al. also found that feed screw rate had minimal effect on P_{max} relative to roll speed and hydraulic roll pressure for a maize starch system.[39] However, maintenance of a constant roll/feed screw speed ratio enabled optimal operating conditions. After reviewing numerous studies, it is apparent that there is no clear "dominant input factor" that governs the roller compaction operation, as the impact of different input factors on product quality is highly dependent upon the material and equipment being used.

One input factor that is consistently assessed in roller compaction studies is the impact of feeding conditions, such as feed rates and feed design, on compaction force (or pressure). This topic of feeding conditions remains of particular interest because of its significant manifestation

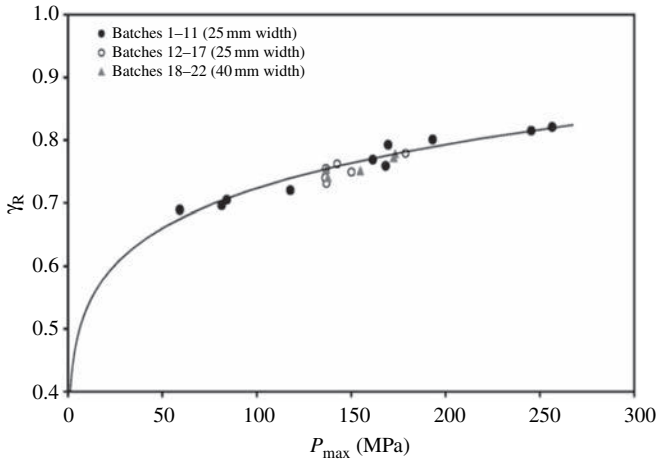


FIGURE 32.14 Ribbon solid fraction (γ_R) as a function of maximum pressure (P_{max}). Roller compacted material is an active blend (>30% drug load) with MCC and dibasic calcium phosphate dihydrate (DCPD) as primary excipients compacted over a wide range of operating conditions and two roller width configurations. Multiple lubricant levels were also explored.[52] *Source: ©2010, with permission from Elsevier.*

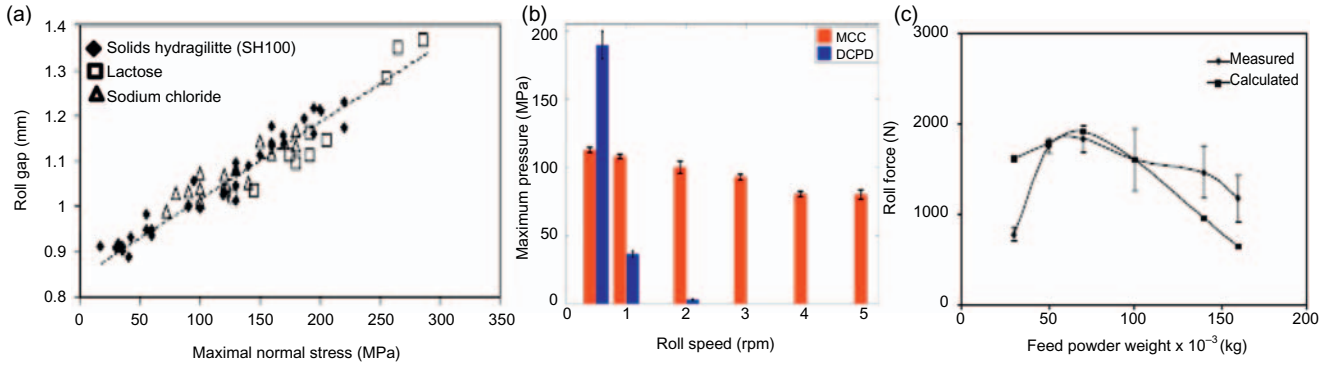


FIGURE 32.15 Impact of roll gap, roll speed, and feed rate on compaction pressure: (a) Roll gap as a function of maximum normal stress for solids hydragilitte, lactose, and sodium chloride,[14] (b) maximum pressure as a function of roll speed for MCC (Avicel PH 102) and DCPD,[50] and (c) roll force as a function of weight of feed fed during the process for maize powder.[55] *Source: ©2003, ©2012, ©2005 with permission from Elsevier.*

on ribbon density. Simon et al. reported heterogeneity in stress patterns along the width of a roller compactor roll and matched its period with that of the single screw feeder in the roller compactor unit. In this work, sodium chloride (NaCl) was roller compacted and the stress experienced by the powder was measured directly using an instrumented roll and the density profiles across the resultant ribbons were quantified using a light transmission technique. The stress exerted on the compact demonstrated a periodical heterogeneity in stress distribution, with patterns matching that of the single feed screw.[17] This phenomenon was also captured for alumina and lactose[17] and was reported by Cunningham when studying stress patterns of roller-compacted MCC.[35] Roller compaction operations that utilize a piston-feeding device, where a uniform pressure delivers powder to the rolls as opposed to the periodical pressure from a rotating screw, did not display periodical stress patterns in the ribbon. Alternatively, Lecompte et al.[15] did not observe periodical stress patterns in roller compacted material even when using a single-screw feed design. In Lecompte et al.'s experimental set-up,[15] the feed screw was positioned further away from the rolls than the unit described in Simon et al.[17] This extra distance between the feed screw and the rolls dampened the screw's ability to impart stress heterogeneously to the ribbon, providing further indication that heterogeneous stress distributions are tied to feeding conditions (ie, screw geometry and position). Along this same principle, stress variations on the powder were shown to be minimized by using a dual-feed screw design in combination with vacuum deaeration.[56] Additional findings from Lecompte et al.'s study found that operating parameters that maximize the amount of powder fed to the rolls (ie, high feed screw speeds, low roll speeds, and narrow gap settings) also serve to promote uniform stress distribution across the roll width.

Interest in understanding heterogeneities in ribbon densities extends beyond a general academic pursuit. Stress heterogeneities in the ribbon can lead to large variations in ribbon densities, which can then impact granule quality in terms of wide particle size distributions and potentially high fines content if the ribbon densities in some regions are especially low. As alluded to previously, multiple studies have reported the existence of pressure gradients across the ribbon width, regardless of feeder type.[16,17,42,57] In Fig. 32.16a, 3D images depict the stress distribution across the width of a ribbon manufactured under varying hydraulic roll pressures. At select operating conditions, where the hydraulic roll pressure was low, there were significant variations in the stress at the center of the ribbon versus the edges of the ribbon, in line with that observed in Fig. 32.16b. As the hydraulic roll pressure was increased, the stress distribution over the width of the ribbon became more homogenous.[15] These stress variations were correlated with density variations across the ribbon width. Ribbon density, as a function of width, was mapped using three different characterization techniques: micro-indentation, X-ray micro-computed tomography, and segmentation. The results from all three techniques were in alignment with instrumented roll data, showing that maximum compaction stress tends to occur near the center of the ribbon.[16] The reason for this density pattern is attributed to frictional resistance on the sides of the rolls, which reduces powder flow to the edges. These stress/density

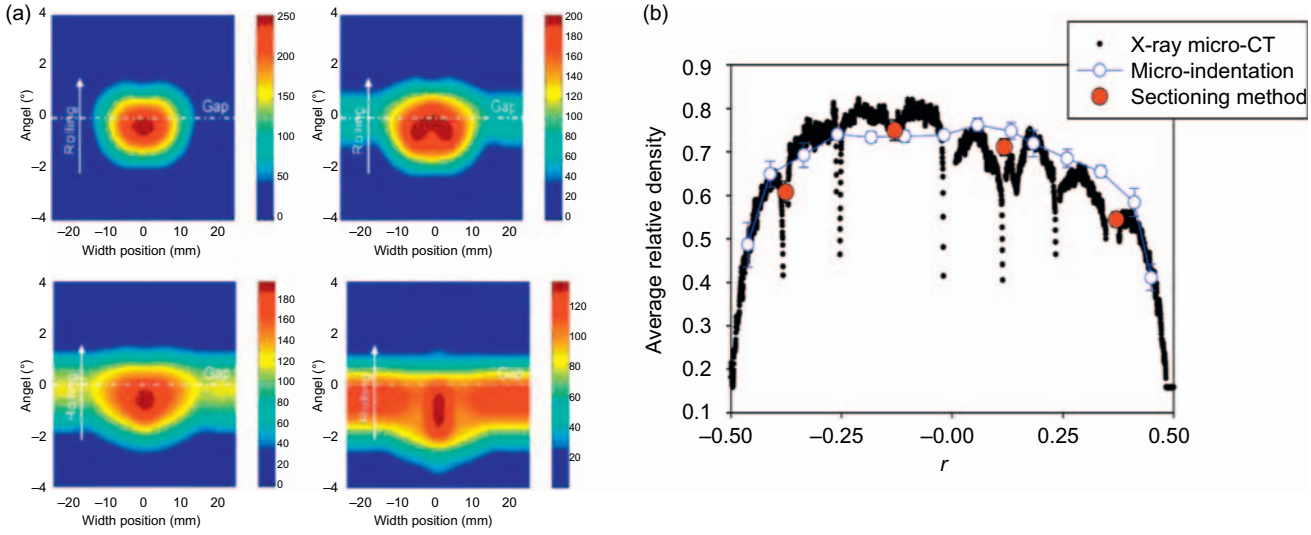


FIGURE 32.16 (a) Stress distributions (MPa) for different compaction forces as a function of rolling angle and roll width position where an increase in hydraulic roll pressure favors more homogenous stress distributions across the roll width,[15] (b) density profile along the ribbon width, where densities have been measured using multiple characterization techniques.[16] *Source: ©2005, ©2009 with permission from Elsevier.*

variations can be minimized by implementing side seals to minimize leakage and imposing a vacuum to facilitate deaeration of the powder.[12,42,56,57] Miguelez-Moran et al. used an instrumented roll to demonstrate that density variations across the ribbon were minimized when roller compactors were run at lower roll speeds, smaller gaps, and when using powders that slipped along the seal surfaces more readily. Lubrication of sealing plates and/or sufficient lubrication of feed powders minimized ribbon density variations, especially those consisting of lower edge densities and higher densities at the ribbon centers, through the reduction of frictional effects between the powder and the sealing plates.[48]

As indicated by Miguelez-Moran et al.,[16] formulation choices can also have significant effects on roller compaction operations. Instrumented roll technology has also been used to explore the impact of formulation considerations, such as filler, binder, and lubricant selections, on roller compaction performance, and more specifically, on ribbon and granule quality. The majority of these studies have focused on the behavior of a specific excipient, or class of excipients, to provide insight into guiding the selection of appropriate materials for roller compaction. In a few cases, a model active ingredient has been added to investigate the impact of the active on excipient behavior. Several studies have examined the role of roller compacting a more plastically deforming material, such as MCC, versus a more brittle, fragmenting material, such as lactose or DCPD. Yu et al. found that a lower P_{max} was required for DCPD to achieve the same solid fraction as MCC when roller compacted under the same conditions.[50] Inghelbrecht et al. conducted extensive studies to understand how roller compaction operating parameters needed to be adjusted to achieve similar quality granules when roller compacting different lactose grades (α-lactose monohydrate 200M, α-lactose monohydrate 90M, anhydrous β-lactose, and spray-dried α-lactose)[58] and different MCC grades (Avicels PH-101, PH-105, PH-302, RC-581, CE-15).[59] The authors concluded that more brittle materials (ie, lactose) required lower feed screw speeds than plastic-deforming materials (ie, MCC) in order to achieve similar quality granules, in terms of friability and granule size. In the MCC study,[59] ibuprofen was added to the blends, up to 75% by weight, and found to disrupt the binding capacity of MCC. For ibuprofen loadings between 25% and 50%, higher compaction pressures (ie, P_{max}) were needed to achieve the same granule quality as pure MCC. However, at 75% ibuprofen content, the brittle nature of the ibuprofen dominated the material's behavior and the target granule quality was attained using lower P_{max} conditions, in line with the earlier observations regarding roller compaction of more brittle excipients (lactose) compared with more plastic excipients (MCC). Interestingly, the extent of the impact of ibuprofen content on granule quality varied for the different MCC grades, which was attributed to differences in particle size and density. The impact of different binder selections, including different grades of copovidones, crospovidones, and celluloses,[60] as well as lubrication levels,[35,48,61] on granule quality have also been reported. As seen in Fig. 32.17, Yu et al. found that the addition of magnesium stearate (MgSt) as a lubricant influenced MCC and DCPD very differently, where increasing levels of MgSt reduced shear stress for DCPD, but MgSt had no effect on MCC. Boundary lubricants, such as MgSt, possess low coefficients of friction, typically ~0.1,[62] and thus provide a weak interfacial layer for materials. Organic materials, such

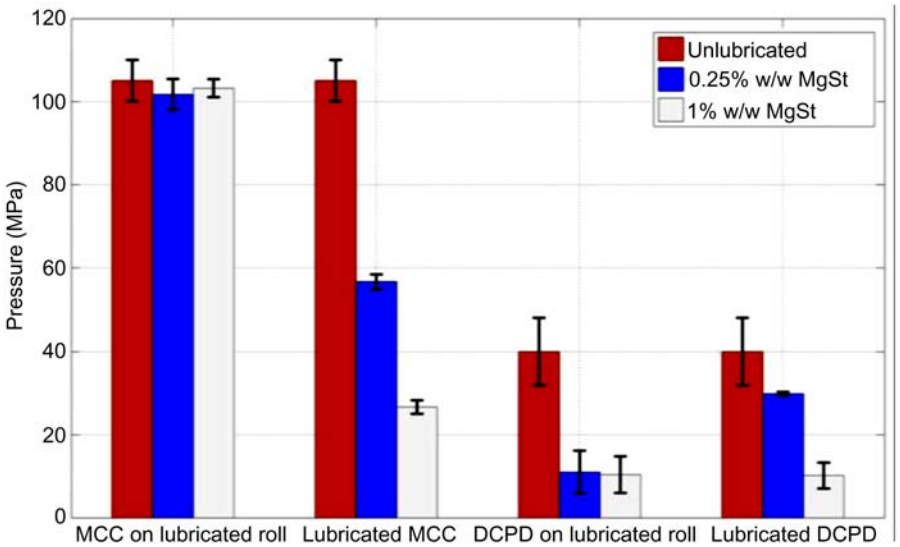


FIGURE 32.17 Impact of lubrication on pressure for MCC and DCPD.[61] *Source: ©2013 with permission from Elsevier.*

as MCC, can be difficult to boundary lubricate because they already possess low interfacial shear strengths, similar to that of boundary lubricants. However, the coefficients of friction for nonorganic materials, such as DCPD, are typically higher than that of organic materials (~0.5 for DCPD), making them amenable to the effects of a boundary lubricant. Wu et al. studied the impact of moisture content of the feed powder on ribbon quality.[63] Results showed that MCC powders containing more than 10% water content were more susceptible to nonhomogenous stress distributions in the ribbon. As the moisture content in MCC powders was increased, the powder became more cohesive and flowability of the moist powders decreased. Consequently, feeding conditions in the roller compactor unit were impacted by nonuniform flow of powders into the compaction zone and exacerbated by increased frictional effects between the powder and side cheek plates. These conditions resulted in P_{max} and density distribution profiles across the ribbon where the values at the center of the ribbon were significantly higher than values at the edges, with higher moisture contents of the feed powder leading to more pronounced differences between center and edge values. Interestingly, ribbons made with powders with high moisture content also possessed different hydration states across the ribbon. In all of these cases, the authors found that optimized roller compaction parameters varied significantly depending on the feed's material properties. Differences in roller compaction behavior were observed when materials possessed different particle sizes, shapes, morphologies, moisture content, and plastic/elastic properties. Overarching conclusions could not be extracted from these studies because many of the identified trends related primarily to excipients, and thus their application to realistic drug formulations has not been fully established.

To summarize, none of the empirical correlations discussed in this section could have been made without the use of an instrumented roll because marketed commercial roller compactors do not have the capability to monitor the actual compaction pressure experienced by the powders. The ability to separate equipment-induced pressure effects from the actual pressure that is directed toward powder compaction is vital for gaining in-depth understanding of the roller compaction process. While ribbon densities are typically monitored to track the adequacy of different roller compaction settings, literature has clearly shown that ribbon densities are influenced by numerous parameters, ranging from material characteristics to process parameters and equipment considerations, and thus, knowledge of actual pressure profiles has been essential to improving the mechanistic understanding of roller compaction behavior. While the instrumented roll has been useful in gaining additional insight into the roller compaction process, it has also revealed the significant complexities of this operation.

32.2.1.2 *Modeling contributions using instrumented roll*

Instrumented roll technology has been a useful tool to support the development of roller compaction models. Schonert et al. used the instrumented roll to substantiate the existence of slip and nonslip compaction regions, a key aspect that serves as the foundation for the majority of roller compaction models.[43] Instrumented roll data has been used to collect input for finite element (FEM)-based models to describe feed stresses induced by the feeder, roll surface pressures, and roll shear stresses at different locations across the roll width.[35,64,65] A more detailed discussion on the use of instrumented rolls for modelling purposes is presented in Section 32.3.

32.2.2 Uniaxial compaction to simulate roller compaction

Uniaxial compaction, using both a compaction simulator and a rotary tablet press, has been used to simulate ribbon manufacture on a roller compactor.[26,49,66,67] Current approaches to feasibility and development, as well as scale-up, of roller compaction operations require significant amounts of material for experimentation. Technical transfer of roller compaction processes from one machine to another, in addition to scale-up challenges, depends on equipment parameters that are prone to bias through equipment type and instrumentation. For example, the accurate translation of maximum compression force exerted by the rolls on the powder requires knowledge regarding powder behavior (eg, flow and compressibility) and knowledge of equipment properties (eg, roll configuration, how force is applied to the rolls, and energy losses due to equipment design). Successful transfer of roller compaction processes typically cannot be confidently assessed until the ribbons have been manufactured and characterized.

A material-sparing approach to the roller compaction process, in which compression events during roller compaction are simulated using uniaxial compaction, has been used to increase the development efficiency of roller compacted products (Fig. 32.18). Limited amounts of material may be used for uniaxial compaction compared with roller compaction (~10× reduction), and characterization of the simulated compacts enables more rapid screening during development.[26] Key process parameters of maximum compaction pressure, roll gap, and material mass can be accurately simulated and controlled; however, it should be

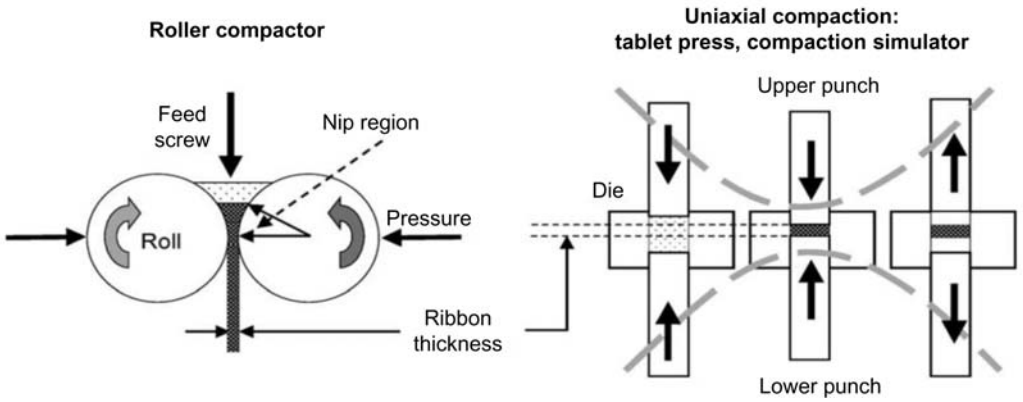


FIGURE 32.18 Simulating a roller compactor using uniaxial compaction.[26] *Source: ©2004 with permission from Elsevier.*

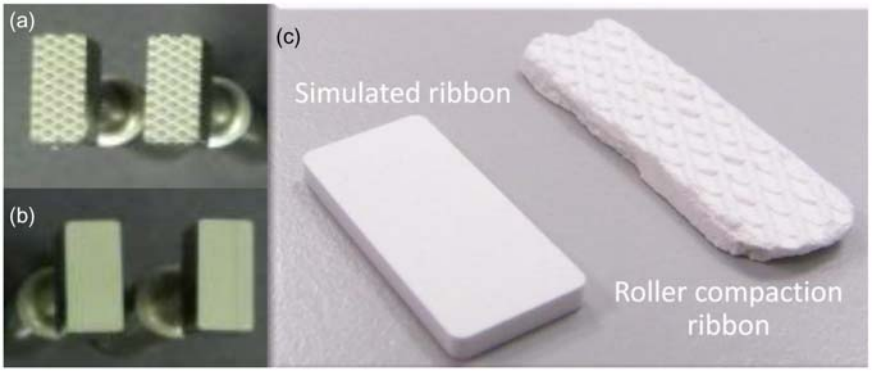


FIGURE 32.19 Rectangular tooling used for roller compaction simulation: (a) smooth surface, (b) knurled surface, (c) photographs of simulated and actual ribbons. *Source: Photographs from unpublished data.*

noted that the simulation approach does not account for variables associated with feed mechanisms and the shear behavior experienced by the powder between the rolls. Ribbon solid fraction is typically used as the primary indicator of product quality for comparison between the simulation using uniaxial compaction and the actual roller compaction process. Uniaxial compaction for roller compaction simulation has been conducted using rectangular-shaped, flat-face tooling in an attempt to mimic the roll surface and to facilitate tensile strength measurements using the three-point bend test (Fig. 32.19), although good correlations between simulated and actual roller compaction processes have been achieved using round flat-faced tablet tooling. In summary, this simulation approach may be used to efficiently determine target ribbon qualities and, if appropriate correlations are developed between compaction pressures and roll pressures, optimal roller compaction operating parameters.

Hein et al.[21] used uniaxial compaction to analyze the suitability of materials for roller compaction. Crystalline α-lactose monohydrate, anhydrous β-lactose, spray-dried lactose, and MCC compacts were assessed to determine which of the diluents would be most suitable for roller compaction. The materials were compacted using an eccentric tablet press equipped with round, flat-faced tooling, milled to form granules using a rotary granulator, and then recompacted on the eccentric press to form tablets. Characterization of resultant tablets (ie, elastic recovery, tensile strength) was compared with previously published studies in which those same materials were roller compacted. The measured properties for these tablets prepared using the simulated roller compaction process compared well with the results found in the literature for tablets manufactured by an actual roller compaction process. Of the materials assessed in this study, anhydrous lactose was found to be the most suitable material for roller compaction, which was in agreement with previous reports.[27,58,59]

Few studies have reported direct comparisons between actual roller compaction experiments and those simulated by uniaxial compaction.[26,48,66,67] Gereg et al.[66] correlated data generated for lactose

monohydrate on a manual hydraulic press (Carver press model C, Carver Inc.) to a production-scale roller compactor (Fitzpatrick, model IR-520), using a roll pressure equivalent to the compaction pressure used on the Carver press, and found that granules (assessed for particle size, flow, and density) and final tablets (characterized by friability, density, and hardness) manufactured by both methods possessed exceptionally similar characteristics. However, ribbon qualities were not assessed in the study. Miguelez-Moran et al.[48] reported that similar pressure-density relationships were obtained for MCC regardless of the method of compaction, uniaxial or roller compaction, and even roll speed, as long as the tablet and the ribbon possessed the same thickness (Fig. 32.20). By keeping the thickness of the compacts constant, the degree of deformation and wall friction effects were similar.

Zinchuk et al.[26] used a custom-built compaction simulator and rectangular-shaped tooling to simulate roller compaction. Unlike experiments using manual hydraulic presses or rotary presses, the compaction simulator was programmed to more closely emulate the roller compaction profile, enabling movement of both top and bottom punches and the ability to program the exact profile of a particular roller compactor machine. In addition, both smooth and serrated punch surfaces were also assessed because the surfaces of rolls are often patterned to facilitate powder uptake in the feed zone. Actual and simulated ribbons of MCC exhibited similar compression behavior (compactability) and mechanical properties (tensile strength), regardless of roll surface pattern, as long as the ribbons were compacted to the same solid fraction (Fig. 32.21). Minor discrepancies that were observed between simulated and actual ribbons were insignificant compared with the variability in the measurement of the actual ribbon attributes, given the nonhomogeneity of densities and strengths across the ribbon width.[17] The extent of these discrepancies between actual and simulated ribbons may depend on material properties. Larger standard deviations were observed for ribbons manufactured at higher compaction pressures and were attributed to differences in micro-crack formation in the compacts at such high pressure conditions. By focusing efforts on maintaining consistent ribbon solid fraction, rather than equipment-dependent factors like compaction pressure, ribbon quality was deconvoluted from instrument type and parameters, and should facilitate consistent process scale-up or transfer between instruments. It was assumed that simulated and actual ribbons possessing equivalent material properties would result in equivalent granulations.

Hilden et al. went one step further and developed a correlation between compaction stresses achieved during uniaxial compaction to that during roller compaction. Based on phenomenological observations, a relationship between peak ribbon stress, σ_{Equiv}, and applied roll force, F, was developed:

$$\sigma_{\text{Equiv}} = 9.1\frac{F}{Rw} \tag{32.1}$$

where R and w are roll radius and width, respectively. This relationship was applied to six different case studies, examining formulations ranging from placebos to active formulations, up to 37% drug load, and multiple roller compactor units (Fitzpatrick IR 22 and IR520, Gerteis Mini-Pactor 250/25). In all cases, the solid fractions of compacts manufactured by

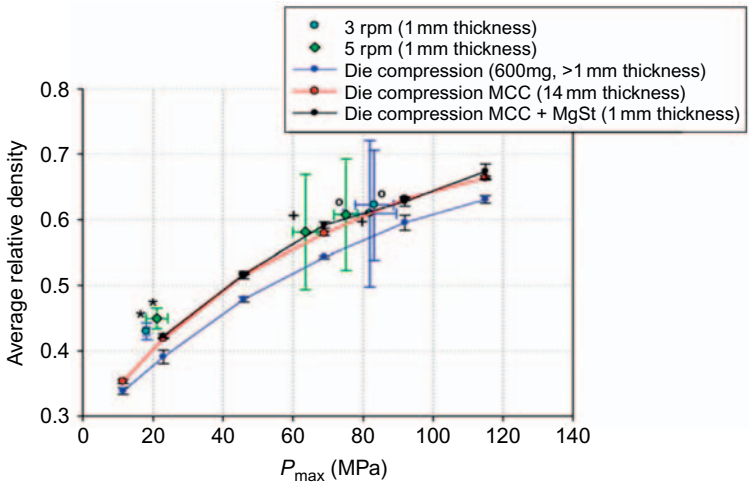


FIGURE 32.20 Average relative density of compact as a function of maximum compression pressure (P_{max}): (+) No lubrication of press or in powder, (o) lubricated press only, and lubricated powder (+).[48] *Source: ©2008 with permission from Elsevier*

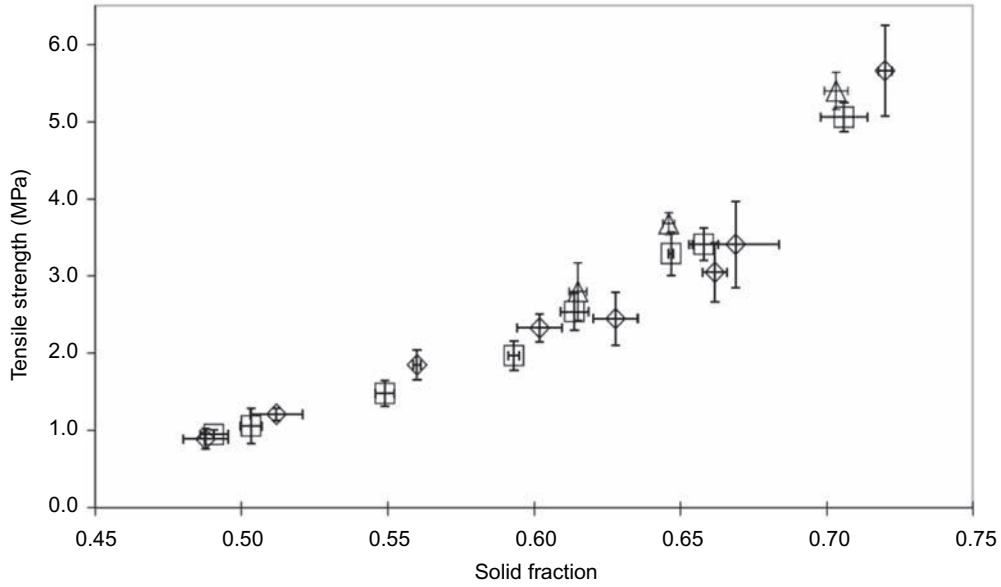


FIGURE 32.21 Tensile strength comparison of actual and simulated ribbons as a function of solid fraction: ◇ smooth real ribbons; □ smooth simulated ribbons; Δ serrated simulated ribbons.[26] *Source: ©2004 with permission from Elsevier.*

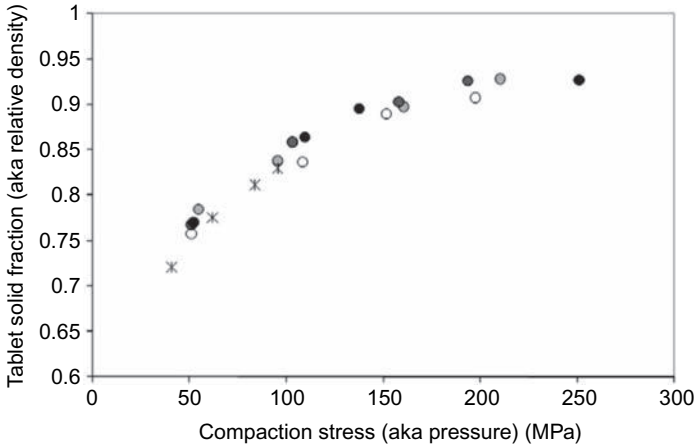


FIGURE 32.22 Overlay of tablet solid fraction as a function of compaction stress. Circles represent data from uniaxially compacted tablets and stars represent data generated on a roller compactor. Different shades of circles denote tablets compressed from ribbons of low (light) to high (dark) solid fractions.[67] *Source: ©2011 with permission from Elsevier.*

uniaxial compaction were comparable to those of actual roller-compacted ribbons when subjected to an equivalent compaction stress. In one example, an active formulation (31% drug load) was roller compacted under multiple conditions to yield ribbons of varying solid fraction. Fig. 32.22 shows how the compaction profiles of different batches of tablets, prepared using the granules of varying solid fractions, lie essentially along the same curve. In addition, the compaction profiles of the roller-compacted ribbons, which did not undergo a tabletting operation, also collapsed upon this same curve. Not only does this data further support the concept of using uniaxial compaction to simulate roller compaction, but it also provides further evidence supporting the concept of a unified compaction curve (described in the introduction section). The value of the relationship between applied roll force and equivalent compaction stress is its ability to provide information regarding intensive material properties and not be convoluted by extensive process or tablet properties.

32.2.3 Density characterization tools

Envelope density is the density used to determine the solid fraction of roller compaction ribbons. Envelope density is defined as the mass of an object

divided by its volume, where the volume includes that of its pores and small cavities. Due to the irregular shape of roller compaction ribbons, determining the envelope densities of the ribbons requires specialized tools, a few of which are listed next.

32.2.3.1 Envelope density analyzer

Envelope density analyzers utilize a displacement technique in which the volume of a solid object, such as roller compaction ribbon, is determined by the displacement of a solid medium. A brief description of the technique is as follows: a sample holder is filled with a dry, solid medium and a plunger compresses the powder until a specified compression force is achieved. This preliminary compaction, containing only the displacement medium, establishes an initial baseline for the volume in the sample holder. Next, the test run is conducted. In the test run, the object is placed in the sample holder containing the medium and the compaction process is repeated. The difference in the distance traveled by the plunger between the preliminary and test run is used to calculate the displacement volume, which can be attributed to the volume of the object. This process, illustrated in Fig. 32.23,[68] is able to accurately capture geometries of patterned ribbons.

The solid medium should be composed of small, rigid spheres that possess a narrow size distribution, good flow properties, and achieve close packing around the object under analysis. The particles of the medium should be sufficiently small such that they conform closely to the object's surface during consolidation, yet do not invade any pore space. The most commonly used commercial envelope density analyzer is the Micromeritics's GeoPyc Model 1360, in conjunction with DryFlo as the medium.

32.2.3.2 Arc punch

An arc punch is utilized in the same manner as a cookie cutter. The punch is pressed upon a ribbon in order to cut out a sample of a known geometry. The mass and thickness of the ribbon sample is measured, thus enabling the calculation of the sample volume. This method does not account for patterned surfaces of ribbons.

32.2.3.3 Laser profilometer

A laser profilometer with dual laser distance sensors (custom designed by Custom Lab Software Systems) was used to determine the density of roller compaction ribbons. The profilometer scans ribbon samples with low intensity laser beams that integrate the sample thickness across the ribbon surface and then quantitate the apparent volume. The use of the laser profilometer method to determine ribbon density was compared with more traditional methods of volume displacement (ie, GeoPyc) and manual measurements using calipers. Fig. 32.24 shows how ribbon density measurements by the laser profilometer compared well with measurements by volume displacement. Advantages of using the laser profilometer over the volume displacement method are more rapid turn-around times and better capability to handle friable ribbons. Both of these methods are significantly more accurate than measurement by calipers.

32.2.4 Ribbon strength characterization

32.2.4.1 Three-point bending flexural test

Tensile strength of roller compaction ribbons may be assessed using a three-point bending flexural test. This technique has been widely used to test mechanical properties of ceramics, polymeric, and metallic materials, and more recently for pharmaceutical powders.[26,70,71] A rectangular sample is stressed by applying a force to the center of the sample until the sample fails (Fig. 32.25). The tensile strength, σ_T, of the compact is determined using the following relationship:

$$\sigma_T = \frac{3}{2}\frac{FS}{Wt^2} \tag{32.2}$$

where F is the load applied to the fracture, W is the sample width, S is the distance between supports holding up the sample, and t is the sample thickness. This relationship assumes that linear elastic behavior

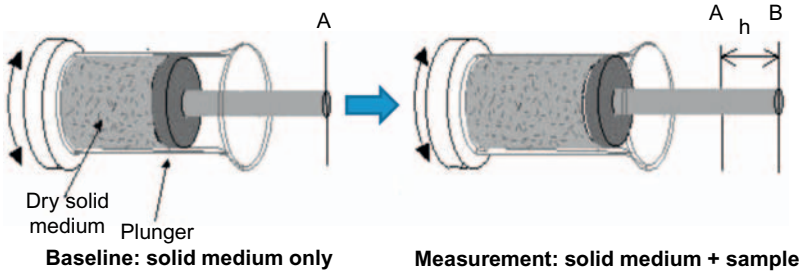


FIGURE 32.23 Envelope density analyzer operation: volume determination by the displacement of a solid medium. *Source: Adapted from Webb PA.* Volume and density definitions and determination methods. *Micromeritics Instrument Corporation; 2001 (Micromeritics Incorporation).*

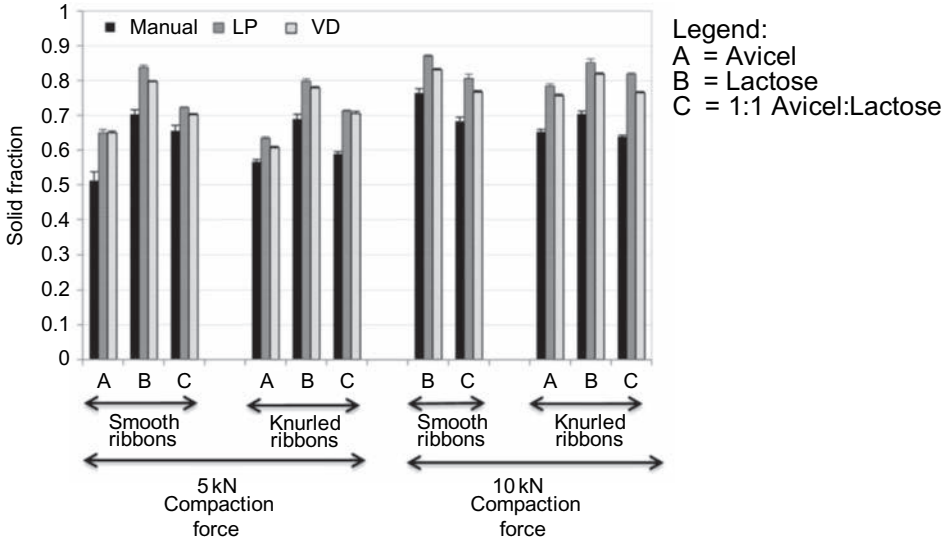


FIGURE 32.24 Solid fraction of ribbons determined using three measurement techniques: *Manual*, caliper measurement; *LP*, laser profilometer; *VD*, volume displacement (GeoPyc). Impact of roll surface pattern and compaction force on ribbon solid fraction is shown. *Source: Adapted from Iyer RM, et al. A novel approach to determine solid fraction using a laser-based direct volume measurement device.* Pharm Dev Technol *2013;**19**(5):577–82.*[69]

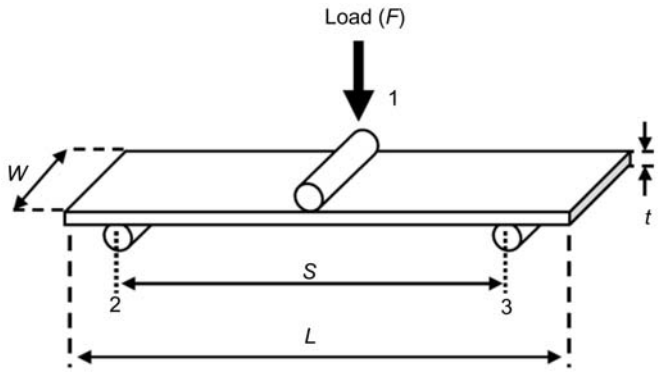


FIGURE 32.25 Three-point bending flexural approach to measure tensile strength of rectangular compacts, where *W*, *L*, and *t* are sample width, length, and thickness, respectively. *S* is the span between sample supports, *F* is the applied load; 1 is the loading point, 2 and 3 are the bottom sample supports.[26] *Source: ©2004 with permission from Elsevier.*

occurs[72] and that stress varies linearly across compact thickness, from maximum tensile stress at the lower face of the compact through zero to equivalent compressive stress on the upper surface of the compact. Tensile strength values were found to be independent of loading rate.[73] To facilitate the direct comparison between samples, compacts tested should optimally possess the same dimensions. Three-point bend tests may be performed using a dynamic mechanical analyzer, universal testing machines, such as an Instron testing system (Instron, High Wycombe, UK), or a texture analyzer.

32.2.5 Process analytical technology tools

Process analytical technology (PAT) tools enable nondestructive, online monitoring of critical process variables, thus facilitating the ability to make real-time adjustments to process operations in order to consistently manufacture products with the desired qualities. PAT tools pave the way for real-time release of products, eliminating the need for extensive offline laboratory testing, and promote the adoption of continuous manufacturing systems. Critical quality attributes (CQAs) for roller compaction include ribbon (density, moisture content, tensile strength) and granule (density, moisture content, tensile strength, particle size) characteristics and content uniformity.[74] Ideally, the objective of implementing PAT tools into the roller compaction operation is to be able to develop and utilize statistical, predictive cause–effect relationship models to implement system controls and feedback loops to automate operations. Fig. 32.26 provides a schematic of how PAT may be used to monitor a roller compaction system.[75] The data collected from PAT tools, in conjunction with multivariate analysis approaches, can be used to elucidate relationships between input parameters and CQAs.

The first form of online monitoring applied to roller compaction was acoustic emission, where the signal produced during powder compression was recorded with a microphone and transformed into a frequency spectrum. Changes to the acoustic emission bands correlated to different compression forces and were

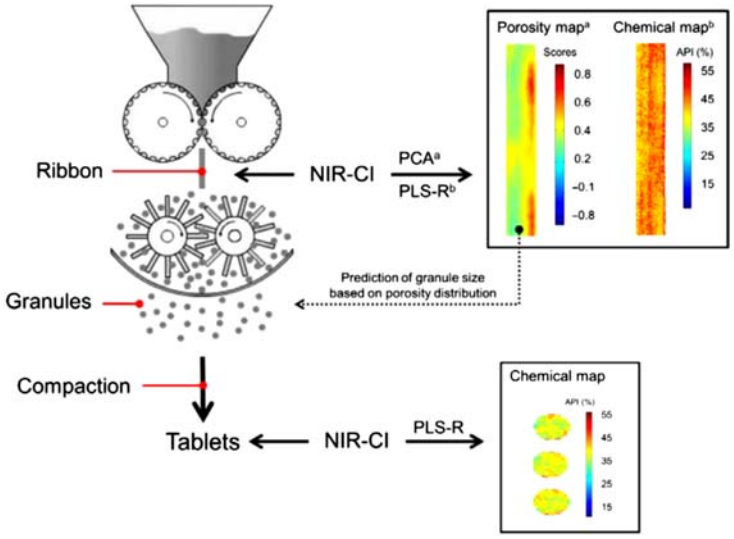


FIGURE 32.26 Overview of process monitoring of roll compaction and tabletting: the implementation of NIR-CI to gain information related to the physical or chemical properties of intermediate or final product.[75] *Source: ©2015 with permission from Elsevier.*

especially apparent when capping occurred. More recently, Akseli et al.[76] demonstrated the use of acoustic wave forms in the ultrasonic spectrum, along with X-ray micro-computed tomography to quantify density variations in ribbons. Heterogeneities in ribbon densities were correlated to tablet mechanical properties to provide insight into how lubrication levels and processing conditions at the roller compaction step impacted subsequent tablet properties.

NIR probes have also been successfully used to monitor ribbon properties, including density,[77–84] moisture content,[78–80] tensile strength,[78,79,83] Young's modulus,[78,79] API concentration,[79,83] and granule properties, including particle size.[77,82,84] NIR data, in conjunction with data analysis methods, have been used to develop models to predict these ribbon and granule properties. Gupta et al.[77] proposed a model to predict ribbon tensile strength based on the slope of the best-fit line through the NIR spectra. Predictive models have also been developed in conjunction with partial least squares (PLS) regression to predict ribbon density, moisture content, tensile strength, and Young's modulus.[78,79] Soh et al.[84] presented a predictive model for granule and ribbon properties that considered not only NIR spectral data, but also raw material properties, operating parameters, and roll gap. Acevedo et al.[81] developed a system to monitor ribbon density in real-time, using not only spectral baseline shifts, but both principal component analysis (PCA) for qualitative analysis and PLS for quantitative analysis. PCA was found to be a more suitable method for monitoring ribbon density because it was a less time-consuming method than PLS. It should be noted that, for the development of NIR-based models, it is important to conduct calibration experiments and subsequent product analysis under similar conditions. Gupta et al. reported large discrepancies between predicted and experimental data when the models were developed using calibration data sets with samples that were stored for 24 hours prior to analysis, whereas experimental data was measured directly from the press.[78] In addition, the selection of optimal probe size and measuring distance (between the probe and the sample) has the potential to yield higher accuracy results.[83]

More recently, microwave resonance tools[80,85] have demonstrated feasibility in monitoring moisture content and envelope density of roller compaction ribbons. Microwave resonance is an attractive option because it does not require sophisticated chemometric software and needs fewer calibration standards than NIR, thus facilitating transfer between measurements of different formulations.[85] These two techniques possess different penetration, being a few millimeters for NIR and several centimeters for microwave; thus microwave signals can penetrate the full thickness of the ribbon. Furthermore, microwave resonance is quite sensitive to water.[74] Consequently, microwave resonance sensors were found to produce more accurate results, compared with NIR (with approximately half of the root mean square error), for both density and moisture determination.[80] However, whereas microwave resonance does not provide chemical information, NIR does.[74]

Near-infrared chemical imaging (NIR-CI), an emerging technique for pharmaceutical applications that is capable of simultaneously collecting large amounts of spectral and spatial information, has recently been applied to develop predictive models for ribbon density, API content, and granule particle size, and can map the density distribution across all ribbon dimensions.[75,86] Using this technique, an NIR spectrum is recorded in each pixel of a sample image, resulting in a

hyperspectral data cube that can be used to visualize the distribution of physical or chemical properties within the sample matrix. Souihi et al. used NIR-CI to effectively determine ribbon density and its distribution across both the width and length of the ribbon, as well as determining the spatial distribution of the compaction pressures on the surface of the ribbon. Results from this study revealed that low roll forces during the roller compaction operation contributed to heterogeneous API distributions in the ribbon. Khorasani et al.[75] used NIR-CI along with PCA to map ribbon density distributions and develop a predictive model for granule particle size distribution. In addition, PLS regression was used to predict API distribution and content within the ribbons. The technology has demonstrated strong potential as a quality control measure for roller compaction operations.

32.3 ROLLER COMPACTION MODELS

Several models have been developed to describe the roller compaction process,[65] such as Johanson's rolling theory for granular solids,[51] the slab method,[4,35,87,88] and modeling by finite element method (FEM)[41,64,89] and discrete element method (DEM).[90,91] The development of these models has not been trivial because numerous complex behaviors must be taken into account[35]:

- Nonlinear behaviors related to the contact or frictional conditions at the roll surface
- Adequate quantification of the constitutive behavior of a deforming powder that undergoes significant evolution of its mechanical properties as a function of density
- Lack of experimental inputs that adequately represent relevant process conditions (eg, frictional coefficient of roll/powder under stress states encountered during roller compaction)
- Experimental difficulties and/or high cost of evaluating the actual process and boundary conditions (ie, feed stress, friction, initial conditions, and density of the powder in the feed region)
- Numerous interrelated parameters associated with equipment design, material properties, and process settings that complicate model verification because many of these parameters can have similar effects on the overall behavior (ie, roll force or density of compact)
- Behavior variability across multiple aspects (ie, rolling direction, across roll width, and even across ribbon thickness) requires two- and three-dimensional effects to be captured, which complicates solution techniques

These process models, in conjunction with experimental studies, have provided important insight into the roller compaction process, enabling a rational scientific framework to probe the underlying mechanisms that drive the operation. This enhanced knowledge has improved our understanding of the relative importance of multiple input factors, including material, equipment, and process considerations, and their impact on product quality. As such, these models play an important role in shaping QbD strategies for the efficient development of roller compacted products and have been used to help establish critical process parameters (CPPs), control strategies, and facilitate scale-up operations. Through the incorporation of models in the QbD approach, a physics-based understanding of the operation is used to predict CPPs, which facilitates a deeper understanding of the compaction profiles experienced by the powder during the roller compaction process that can be used to inform on technical transfer and scale-up activities. For instance, model outputs can be used to identify input parameters necessary to match compaction profiles across the ribbon for different roller compaction units. It should be emphasized, however, that the roller compaction models presented here have been tested only under limited conditions and require additional testing for universal applicability. A summary of process models that have been developed for roller compaction are listed in Table 32.4. To date, roller compaction models have focused primarily on establishing a relationship between input variables (material properties, process conditions) and ribbon properties (solid fraction, ribbon strength). It is commonly assumed that similar ribbon properties lead to comparable granule properties (compactability, flow), which are of interest in downstream operations.

32.3.1 Johanson's rolling theory for granular solids

Johanson's theory is a one-dimensional, continuum-level approach to model the roller compaction process that has been widely used since the 1960s because it has yielded relatively good correlation with experimental data. This theory assumes that powders are isotropic, frictional, cohesive, compressible, and behave in accordance with the effective yield functions proposed by Jenike and Shield.[51,106] As such, the plane strain condition between the rolls can be represented as in Fig. 32.27.[51] Johanson's theory centers upon the existence of three regions of material behavior: a slip, nip, and release region. A detailed description of each of these compaction regions is given in the legend to Fig. 32.28. Briefly, in the slip region, the powder experiences "slipping" at the roll surface and the roll moves

TABLE 32.4 Summary of Process Models for Roller Compaction

Model/reference	Model type	Notes
Johanson[51]	1D physical	Designation of slip and nip regions to model powder behavior analytically
Reynolds et al.[52]; Rowe et al.[53]; Esnault et al.[92]; Nesarikar et al.[42,57]	1D physical	Modified Johanson's theory, which was modified to incorporate elements of feed conditions and parameters; Esnault et al. includes air entrapment influences
Katashinskii[93]; Katashinskii and Shtern[94,95]; Lee and Schwartz[96]; Musikuin[97]	1D physical	Slab method based on constitutive model
Shima and Yamada[87,98]	1D physical	Slab method based on energy continuum and using upper bound theorem
Dec and Komarek[99,100]	1D physical	Slab method using the "plastic" Poisson's ratio
Dec et al.[65]; Cunningham[35]; Cunningham et al.[64]; Michrafy et al.[41]; Muliadi et al.[89]; Garner et al.[101]	2D physical	FEM using modified Drucker–Prager/cap (DP/C) model
Cunningham[35]; Cunningham et al.[64]; Michrafy et al.[102]; Muliadi et al.[103]	3D physical	FEM using modified DP/C model
Odagi et al.[91]	2D physical	DEM; model did not yield satisfactory results
Inghelbrecht et al.[39]; Turkoglu et al.[104]	Artificial intelligence	Neural net
Mansa et al.[105]	Artificial intelligence	Combined neural net and fuzzy logic software package

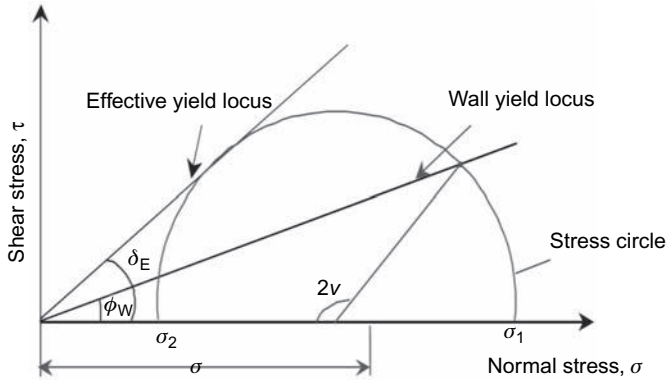


FIGURE 32.27 Jenike–Shield yield criterion for the slip region.[40]
Source: ©2005, with permission from Elsevier.

faster than the powder. At this stage, the material does not undergo significant compaction and the particles rearrange as the rolls "slip" past the material, adhering the Jenike–Shield yield criterion. In the nip region, the material becomes increasingly trapped between the rolls, moving at the same speed as the roll, and undergoes significant compaction. As such, the powder no longer follows the Jenike–Shield yield criterion and behaves according to a powder bulk density-hydrostatic pressure relationship. The nip angle, α, is defined as the point at which the material transitions from the slip (a slip boundary condition) to the nip (no-slip boundary condition) region. By convention, an angle of 0 degree corresponds to the point at which the roll gap is at its minimum. A schematic of the roller compaction process is illustrated in Fig. 32.28.

The pressure gradient ($\partial\sigma/\partial x$) in the slip region is described by the following equation:

$$\frac{1}{\sigma}\left(\frac{\partial\sigma}{\partial x}\right)_{\text{slip}} = \frac{4\left[\frac{\pi}{2}-\theta-\upsilon\right]\tan(\delta)}{\frac{D}{2}\left(1+\frac{S}{D}-\cos\theta\right)(\cot(A-\mu)-\cot(A+\mu))} \tag{32.3}$$

where δ is the feed material's effective internal angle of friction, D is the roll diameter, S is the roll gap, φ_w is the angle of friction between the roll face and the feed material, μ is the coefficient of friction between the powder and roll surface, and A is defined as:

$$A = \frac{1}{2}\left(\theta+\upsilon+\frac{\pi}{2}\right)$$

$$\upsilon = \frac{1}{2}\left[\pi-\sin^{-1}\left(\frac{\sin\varphi_w}{\sin\delta}\right)-\varphi_w\right]$$

The pressure gradient in the nip region is described by:

$$\frac{1}{\sigma}\left(\frac{\partial\sigma}{\partial x}\right)_{\text{nip}} = \frac{K\left[2\cos\theta-1-\frac{S}{D}\right]\tan\theta}{\frac{D}{2}\left(1+\frac{S}{D}-\cos\theta\right)\cos\theta} \tag{32.4}$$

where K is the compressibility constant of the feed material.

A summary of the input parameters required by Johanson's theory, as well as output obtained, is listed in Table 32.5.

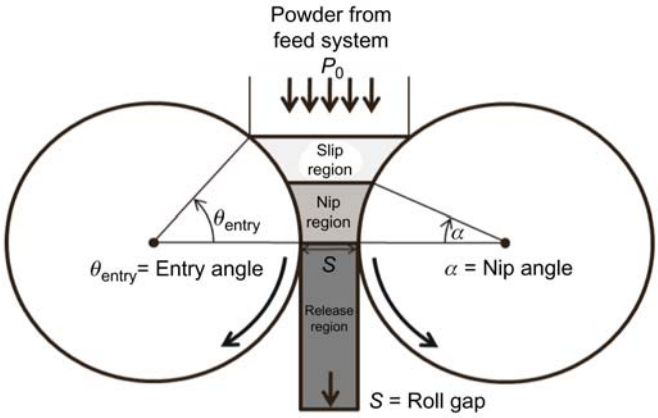


FIGURE 32.28 Schematic of the roller compaction system, where P_0 is the feed pressure, θ_0 is the angle at which P_0 is applied, α is the nip angle, and S is the roll gap.
Slip Region—Powder enters into the area between the rolls by the forward velocity generated by the feeding mechanism (screw feeder pictured in Fig. 32.2). The slip region begins at the point where the rolls begin to impact powder behavior, the boundary of which is defined by the entry angle (θ_{entry}) and the feed pressure (P_0). In this region, the roll surface moves faster than the powder; therefore the powder "slips" on the surface of the roll and compaction forces are very weak. Particle rearrangement in this region is prevalent, facilitating deaeration of the powder. The transition point from the slip region to the subsequent nip region is influenced by a number of factors, including the coefficient of friction between the powder and roll surface, the internal friction coefficient of the powder, design of the roll surface (eg, smooth or knurled), and roll diameter. Powders with low bulk density benefit from longer residence times in the slip region, which can be achieved by lower feed rates and roll speeds, because it enables more time for deaeration and powder densification.
Nip Region—(compaction region) Powder moves at the same speed as the roll surface and is drawn into the roll gap. The boundary between the slip and nip regions is defined as the nip angle (α). Compaction of the powder occurs in the nip region, with the maximum pressure applied just before the point that the roll gap (S) is smallest. Increased residence time in the nip region, due to a large nip angle or a slow roll speed, allows for prolonged application of pressure to the powder. High residence times in the nip region tend to lead to greater particle deformation, breakage, and bond formation, thus producing a compact of greater tensile strength.
Release Region—(extrusion region) Following the nip region, the powder, now in the form of a ribbon, separates from the rolls and transitions into the release region. Elastic recovery and expansion occurs. *Source: Adapted from Rowe JM, et al. Mechanistic insights into the scale-up of the roller compaction process: a practical and dimensionless approach. J Pharm Sci, 2013;102(10):3586–95.*

TABLE 32.5 Input Parameters Required by Johanson's Roller Compaction Model and Output Generated

Input parameter	Experimental measurement technique	Output
MATERIAL PROPERTIES		
Effective angle of internal friction (δ)	Shear test	Normal stress distribution
Wall friction angle (φ_w)	Shear test	Roll force
Compressibility factor of material (K)	Uniaxial compaction	Roll torque
Roll diameter (D) and width (W)	Equipment specification	Nip angle
Feed pressure (P_0)	Roll instrumented with pressure sensors	Ribbon density at $\theta = 0$ degree

At the nip angle, the pressure gradients in the slip and nip regions were proposed to be equal, which allows for the determination of the nip angle by equating Eqs. (32.3) and (32.4). Once the nip angle has been established, the pressure, or stress, profiles of the rolls can be determined if the initial condition is known (P_0). Following a simple power law relationship, the resultant ribbon density (γ) can be related to pressure (σ) as follows:

$$\frac{\sigma_1}{\sigma_2} = \left(\frac{\gamma_1}{\gamma_2}\right)^K \tag{32.5}$$

As such, output from Johanson's model enables an estimation for ribbon density for a specific material/processing condition combination. Alternatively, once the model is established, it can be used to reverse-calculate necessary input parameters to achieve target ribbon densities.

Criticisms of Johanson's theory concern the little attention given to feeding conditions, which are known to impact compaction properties, not accounting for throughput rates (ie, feed screw speed, roll speed), as well as key assumptions regarding the model's stress calculations.[107] However, its simplicity and ability to be solved analytically provides a good basis for modeling efforts, and reasonable agreement between predicted and experimentally measured roll pressures have been demonstrated in studies with MCC (Avicel PH102)[40] and maize powder[55] under limited conditions. In both studies, the predicted results were found to be sensitive to

model input parameters, especially those relating to the feed powder, specifically friction coefficients (δ and φ_w) and feed pressure (P_0).

Difficulties in determining accurate values for the friction coefficients and feed pressure have limited the practical application of Johanson's theory. The shear test is a common experimental method used to determine the friction coefficients required by Johanson's model. However, the accuracy of using such tests has been disputed because the actual frictional behavior of powders (and thus the coefficients) changes significantly over the course of powder densification. Rowe et al. determined the friction coefficient values using a shear tester (Freeman Technology FT4) with varying operating conditions, using consolidation pressures between 9 and 75 kPa. Relatively little change (<20%, ±5 degree) in δ and φ_w values was observed over this range of consolidation pressures.[53] The sensitivity of a modified Johanson-based model was tested over these ranges of δ and φ_w values and was found to have only a minor impact on predicted outputs, which is in agreement with the more exhaustive study conducted by Reynolds et al.[52] using a ring shear tester that examined φ_w changes over a 20 degree^o span and δ changes over a 50 degree span with minimal impact to modeling results. The primary hurdle, however, to the ubiquitous application of Johanson's theory is most likely due to the difficulty in determining P_0 without sophisticated pressure sensor-instrumented rolls (which will be further discussed in this section). P_0 encompasses several important factors, including the extent of material densification in the feeding zone, prior to the slip region, and how efficiently the material passes through the rolls (eg, material jamming prior to the rolls, side leakage). Thus, the value of P_0 significantly impacts modeling results, and has essentially served as a calibration factor that requires adjustment for each feed material-equipment combination. To date, the use of instrumented rolls is the only way to experimentally determine P_0.

The instrumented roll has also been an especially useful tool to address controversies over the proper determination of nip angle. Some have estimated nip angle by determining the angular position at which compaction pressures just begin to increase, but before P_{max} is achieved. Bindhumadhaven et al. estimated nip angle values directly from instrumented roll data through the determination of the intersection of two tangents drawn through the pressure profile, as illustrated in Fig. 32.29a.[40] Feed pressure values were estimated from instrumented roll data. Others have preferred to estimate nip angle using Johanson's model (described in further detail in the next section). For example, instrumented roll data has been used to fit Johanson's stress gradient equations (Eq. (32.3) and (32.4)) to solve for nip angle.[50] Friction angle and compressibility factor parameters (δ, φ_w, K) were determined by multivariate fitting, thus eliminating the uncertainty surrounding the accuracy of offline

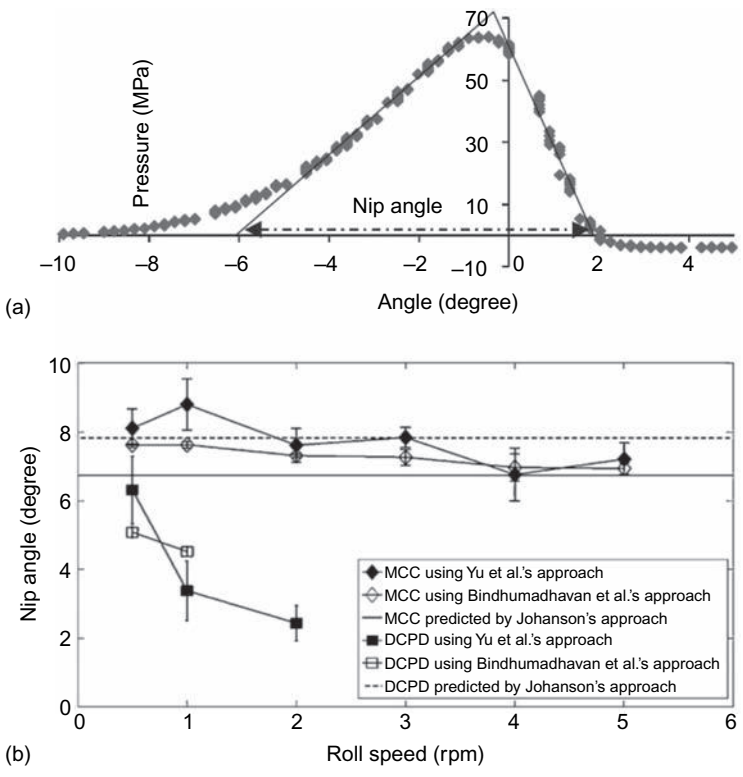


FIGURE 32.29 (a) Nip angle determination from instrumented roll data,[40] (b) variation of nip angle with roll speed for MCC (Avicel PH 102) and DCPD.[50] *Source: ©2005, ©2012, with permission from Elsevier.*

techniques typically used to measure these parameters. Fig. 32.29b compares nip angle determination by multiple approaches.[40,50,51] Predicted nip angle values for plastic-behaving MCC were comparable for all approaches, although Johanson's model clearly did not account for roll speed. However, in the case of DCPD, for which nip angle was more sensitive to roll speed, Johanson's model significantly overestimated nip angle. The discrepancies between Johanson's model and the other two models, especially at high roll speeds, were attributed to inaccuracies in the proper determination of feed pressure values, especially as the feed pressures for poorly flowing materials are often impacted by roll speed and subsequent air entrapment influences. In addition, Johanson's model assumes sufficiently dense packing of powders in the slip zone that facilitates flow of feed powder to the rolls, which is not necessarily representative behavior for poorly flowing materials and may lead to inaccurate estimations of frictional behavior. Bindhumadhaven et al.[40] and Yu et al.'s[50] approaches to nip angle determination again produced comparable results, even for DCPD. Advantages to Yu et al.'s[50] approach is its applicability to low roll speed/small P_{max} regimes, for which Bindhumadhaven et al.'s[40] method is limited, and the elimination of direct measurements of friction angle and compressibility factor parameters, which are intrinsically difficult to measure under appropriate conditions.

Nesarikar et al.[42] has also used instrumented roll data to fit Johanson's roll force equations[51] to estimate nip angles directly from experimental pressure profile data, circumventing the need for input feed pressure values. One of the key features of this model was that wall friction angle values, φ_w, were also estimated indirectly from instrumented roll data, enabling a more accurate representation of frictional behavior between the powder and rolls. Johanson's model, used in conjunction with optimal normal stress profiles for a placebo calibration data set, was found to successfully predict ribbon densities of several active formulations. In addition, by matching the normal stress profiles across two scales of operation, the model was shown to adequately predict scale-up parameters from a pilot (roll diameter: 120 mm, roll width 40 mm) to a commercial-scale (roll diameter: 200 mm, roll width 75 mm) roller compactor. In a separate study, Nesarikar et al.[57] further leveraged instrumented roll data to develop a statistical model, built upon normal stress data collected using instrumented roll technology, that predicted ribbon densities for roller compaction processes. By correlating normal stress data of new materials to that of a previously measured calibration set (based on placebo data), an expression for ribbon density as a function of P_{max} and roll gap was developed. Even though the model was built using placebo data, predicted ribbon densities for placebo and active formulations (up to 17% drug load) compared well with experimentally measured density values. The model was found to be sensitive to varying formulation (lubrication, active drug loadings) and process (roll speed, screw speed, vacuum level, and associated deaeration) conditions. When using statistical models, caution should be taken if using the model outside of the conditions from which the model was built. In this study by Nesarikar et al., %RSD of the prediction increased with increasing drug load.

32.3.2 Modified approaches to Johanson's theory

In 2010, Reynolds et al. presented a modified approach to Johanson's theory that incorporated additional processing parameters to account for material behavior, such as potential precompression of the powder in the feed and slip region, and material throughput during roller compaction.[52] The preconsolidation relative density, γ_0, may be obtained using the following relationship, generated from a material's compaction profile:

$$\log(\gamma_R) = \frac{1}{K}\log(\sigma) + \log(\gamma_0) \tag{32.6}$$

where γ_R is the relative density of the ribbon, K is the compressibility factor, and σ is the pressure. This preconsolidation factor provides additional information regarding a material's compressibility, specifically under low amounts of compaction pressure.

An additional modification to Johanson's theory includes the incorporation of the feed screw speed, N_s, as many roller compactor designs utilize a screw conveyor to transport the material from the feed hopper to the rolls. Typically, a roller compaction operation maintains either a constant roll gap (gap-control) or a constant feed screw speed (screw-control), depending on the formulation. Johanson's original theory only accounts for roll gap. In order to incorporate N_s into Johanson's theory, a mass balance was set up around the nip region, equating the material being fed into the rolls to that exiting the rolls as a ribbon, assuming a steady-state condition with negligible side-seal leakage. A depiction of the mass balance around the nip region is illustrated in Fig. 32.30. The mass flow rate of material entering the rolls is assumed to be proportional to N_s and the material exiting the rolls is a function of the roll speed (N_R) and the solid fraction of the ribbons, γ_R. Based on the mass balance around the nip region, S may be related to N_s and N_R using the following expression:

$$\frac{S}{D} = \frac{c_s}{\pi\rho_{true}\gamma_R D^2 W}\frac{N_S}{N_R} \tag{32.7}$$

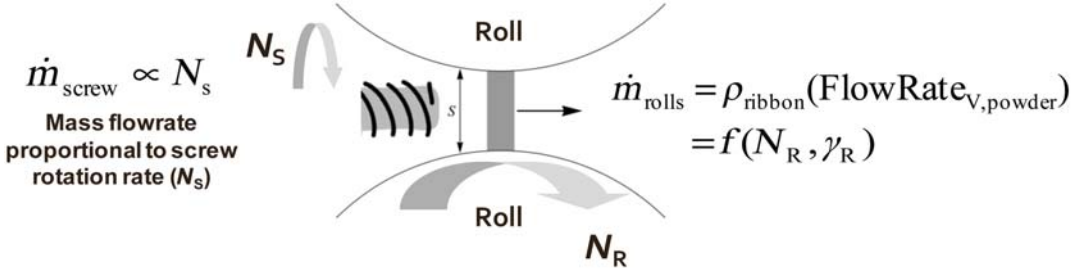


FIGURE 32.30 Schematic of the compaction zone of a roller compactor, where a mass balance is applied to the nip region of the roller compaction process.[53] *Source: ©2013 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.*

where c_s is the screw speed constant, ρ_{true} is the true density of the powder, and W is the roll width. The factor c_s is determined using a set of calibration roller compaction experiments and is used to infer how efficiently the feed screw transfers material into the rolls. This expression for S/D, presented in terms of N_S/N_R, may then be substituted into the equations of Johanson's theory (Eqs. (32.3)–(32.4)) to yield the following expression for the ribbon solid fraction, γ_R

$$\gamma_R = \gamma_0(2R_f)^{1/K}\left\{ WD\int_{\theta=0}^{\theta=\alpha} \times\left[\frac{c_S N_S}{\pi\rho_{true}\gamma_R D^2 W N_R\left(1+\frac{c_S N_S}{\pi\rho_{true}\gamma_R D^2 W N_R}-\cos\theta\right)\cos\theta}\right]^K \cos\theta\partial\theta\right\}^{-1/K} \tag{32.8}$$

where R_f is the roll separating force originally defined by Johanson.[51] The incorporation of γ_0 into the calculation of γ_R was achieved by correlating the final ribbon density to the material's preconsolidated state through the power-law relationship presented by Johanson's theory in Eq. (32.5).

$$P_{max} = \left(\frac{\gamma_R}{\gamma_0}\right)^K \tag{32.9}$$

Reynolds et al.[52] demonstrated good agreement between predicted and experimentally measured ribbon densities across two different scales of roller compactor (25- and 40-mm roll widths), while adjusting common operating factors such as N_S, N_R, S, and hydraulic roll pressure. A key advantage of the approach by Reynolds et al.[52] is the limited amount of material required to build the model, about 1–2 kg, which is used to determine the c_s value. A summary of the input parameters required by Reynolds et al., which are in addition to those required by Johanson's theory (except for P_0), is listed in Table 32.6. The material assessed in this work was an active blend with a drug load greater than 30% wt with primary excipients being MCC and DCPD.

TABLE 32.6 Input Parameters Required by Reynolds et al. That Are Additional to Those Required by Johanson's Theory, Which Are Listed in Table 32.5[52]

Input parameter	Measurement technique
Preconsolidation relative density (γ_0)	Uniaxial compaction (same from K determination)
True density (ρ_{true})	Helium pycnometry
Screw speed constant (c_s)	4–5 roller compaction batches, varying N_s/N_R, and envelope density analyzer to measure ribbon densities

Rowe et al.[53] elaborated upon the work reported by Reynolds et al.[52] to develop a roller compaction model more accessible to formulators and to present a dimensionless approach to modeling the roller compaction process. As mentioned previously, one of the primary deterrents to the universal application of Johanson's model is the difficulty in accurately determining P_0 values without the use of specialized equipment.[51] Reynolds et al.[52] did not discuss the determination of P_0 because roll force values (R_f) were obtained using an experimentally determined calibration factor to relate hydraulic roll pressure to R_f. This calibration factor is not always known, drifts over the lifetime of the equipment, and does not translate between roller compactor units. Rowe et al.[53] proposed an iterative algorithm to determine P_0, for which a series of equations are solved to determine the P_0 value that satisfies both Johanson's theory[51] and Reynolds's modifications,[52] thus eliminating the need to experimentally determine P_0 values. Like the Reynolds's method,[52] the proposed method requires only standard, routinely measured parameters as inputs and minimal amounts of material (~1–2 kg) to establish critical relationships between material input and process parameters. Good correlations between

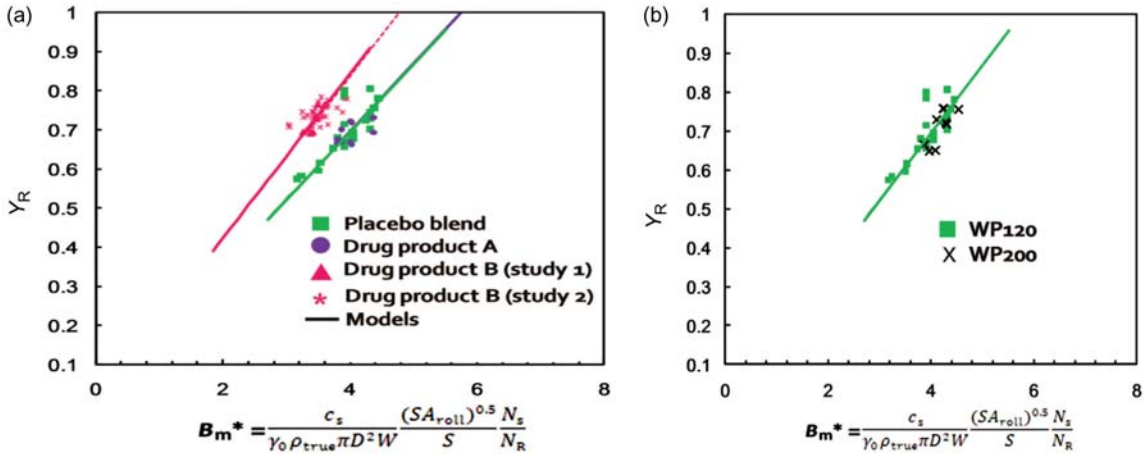


FIGURE 32.31 Solid fractions (experimental data = points, model = line) of ribbons prepared by roller compaction, as a function of a modified Bingham number, B_m^*: (a) ribbons manufactured using an Alexanderwerk WP120 roller compactor, (b) comparison of placebo ribbons prepared on an Alexanderwerk WP120 and WP200.[53] *Source: ©2013 by John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.*

predicted and experimental results were demonstrated for placebo and active blends (up to 22% wt drug load).

In addition to facilitating the mainstream use of Johanson's model, a dimensionless relationship between key process parameters and final compact properties was elucidated to improve the mechanistic understanding of the roller compaction process. This dimensionless parameter, referred to as the modified-Bingham number (B_m^*), relates yield-to-viscous stresses during roller compaction and delineates how operating conditions impact ribbon properties. Yield and shear stresses essentially drive the roller compaction process. Too large of a yield stress can adversely impact the efficiency of material flow through the rolls and will produce a very hard, brittle ribbon that may not recompress well into tablets. Too large of a viscous stress indicates that the material is flowing well through the rolls, but experiencing too little powder densification, producing very soft ribbons that result in poor granules. B_m^* was derived from the fundamental definition of shear stress, and is described as

$$B_m^* = \frac{c_s}{\gamma_0 \rho_{true} \pi D^2 W} \frac{(SA_{roll})^{0.5}}{S} \frac{N_s}{N_R} \tag{32.10}$$

where, c_s is the screw speed constant, SA_{roll} is the surface area of the roll, N_s is feed screw speed, γ_0 is the preconsolidation factor, ρ_{true} is the true density of the powder, D is the roll diameter, W is the roll width, S is the roll gap, and N_R is the roll speed. A linear relationship between γ_R and B_m^* was observed for all of the materials studied, demonstrating the significant correlation between the dimensionless parameter relating yield and viscous stresses to the roller compactor product (Fig. 32.31). The study showed that a constant ratio of yield-to-viscous stresses, as indicated by a constant B_m^*, enabled the manufacture of products with consistent attributes, even at two scales of operation. B_m^* was shown to provide guidance toward determining the design space for formulation development, as well as to facilitate scale-up development.

As mentioned in previous sections, permeating air in the powder bed is known to have a negative impact on the roller compaction process, as the feed powder can be destabilized by the flow of escaping air. Esnault et al.[92] proposed a model to capture these effects, using Johanson's theory to capture solid stress and deformation and Darcy's law for fluid transport. The Kozeny–Carman law was used to describe permeability as a function of porosity and particle size. For this model, the effect of air pressure on the solid was considered to be negligible between rolls and airflow through the gap was ignored. The model predictions were not validated experimentally.

32.3.2.1 Johanson's model comparisons with experimental data (instrumented roll)

A few studies have attempted to quantify Johanson's model predictions against experimental data, which has been enabled by the use of instrumented roll technology. Yusof et al. demonstrated reasonable agreement between measured roll force and roll torque values (using instrumented roll data) and those predicted by Johanson's model for the roller compaction of maize flour as long as roll gap was less than 0.15 mm. Discrepancies between experimental and predicted results for larger roll gaps were attributed to the high sensitivity of Johanson's model to input parameters, especially regarding accurate characterization of frictional behavior. Bindhumadhaven

et al.[40] found that while the Johanson model predicted P_{max} value well for low roll speeds, discrepancies surfaced when roller compaction was conducted at higher roll speeds (Fig. 32.32a). The inconsistencies between experimental and predicted values at higher roll speeds were attributed to the inability of the model to capture the decreasing nip pressures and greater entrainment of air at higher roll speeds. Studies have also shown weaknesses in Johanson's model when material behaviors were not ideal. In Yu et al.,[50] Johanson's theory was shown to overestimate nip angles for DCPD and DCPC-MCC mixtures with <50% MCC due to poor flowability of the feed powders (Fig. 32.32b). To mitigate these gaps in Johanson's theory, instrumented roll data may be used to fit input parameters to Johanson's equations, thus using experimental roller compaction runs to empirically determine the appropriate values for input parameters (ie, P_0 and friction coefficients). Once the appropriate input values have been defined, Johanson's model can be used to describe future roller compaction behavior of the material. Excellent correlation between experimental and predicted results has been achieved by this method.[42,57] In fact, by fitting roller compaction data to Johanson's equation using experimentally determined P_0 values, the pressure profiles of the roller compaction process are elucidated and have been shown to facilitate scale-up activities.[42,57]

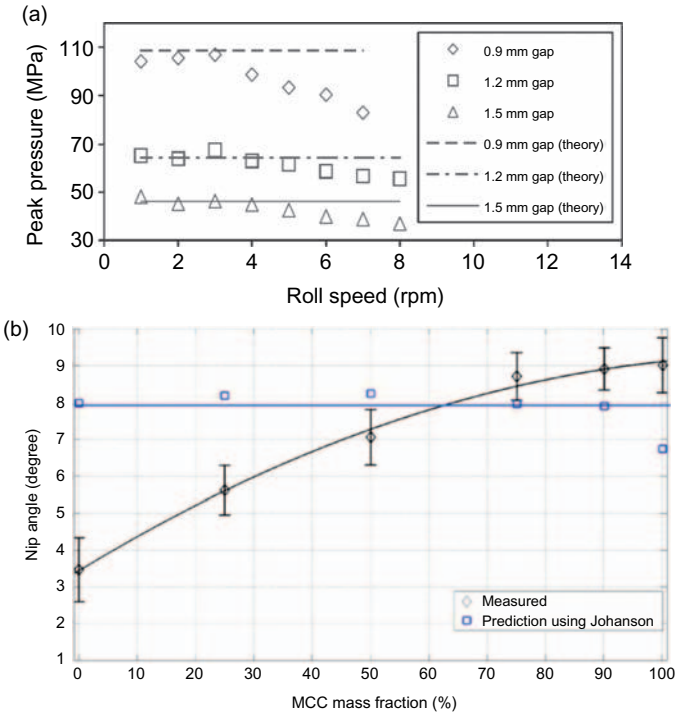


FIGURE 32.32 (a) Peak pressure values as a function of roll speed for different roll gaps. Curves indicate simulated results, as determined from Johanson's theory.[40] (b) Comparison of measured and predicted nip angles for MCC-DCPD mixtures, as a function of MCC content.[50] *Source: ©2005, ©2012, with permission from Elsevier.*

32.3.3 Slab analysis

One of the most fundamental approaches used to model the roller compaction process is the slab method. The slab method considers equilibrium force and mass balance equations on one-dimensional elements, or "slabs," that represent deforming powder, providing information about two principal stresses during compaction, as opposed to Johanson's model which only provides the mean stress. In addition, it allows users to incorporate complex powder qualities to the model, such as cohesion and global friction between particles, making it an attractive modeling option compared with Johanson's approach.

In the slab method, the compaction zone is segmented into differentially small slices, or slabs, and a one-dimensional analysis of the stresses acting on each of these slabs is used to simulate a single element passing through the rolls of a roller compactor (Fig. 32.33). The shape of the slab is bounded by the roll surfaces and, due to the small width (dx) of the slab relative to its height (dh), the boundaries may be represented by straight lines tangent to the roll's contours. Inertial forces acting on the material are considered to be negligible and thus, only the normal and shear stresses acting on the four sides of the trapezoidal slab are assessed. Shear stresses in the system originate from the frictional forces existing between the material and roll surface and the dissection of these stresses into their direct and tangential components enable the calculation of the forces acting on the slab. Equilibration of the forces acting on the roll (along the x direction) can be described by the following equation, which is referred to as the von Karman equation of rolling, and is independent of material properties[108]:

$$\frac{\partial(\sigma_x h)}{\partial\theta} + R(P\sin\theta \pm \tau\cos\theta) = 0 \qquad (32.11)$$

where R is the roll radius, h is the thickness, or height, of the slab, and "+" is applied when $\theta > \theta_{neutral}$ and "−" is applied when $\theta < \theta_{neutral}$. The neutral angle lies between the nip angle and 0 degree, which is the point at which the direction of the frictional force on the roll reverses direction. For instance, when $\theta > \theta_{neutral}$, frictional forces direct powder toward the gap. When $\theta > \theta_{neutral}$, the frictional forces direct powder away from the gap.

The mass conservation and momentum relationships are linked to one another using empirically determined stress-strain constitutive relationships, resulting in a system of equations that describes the normal roll

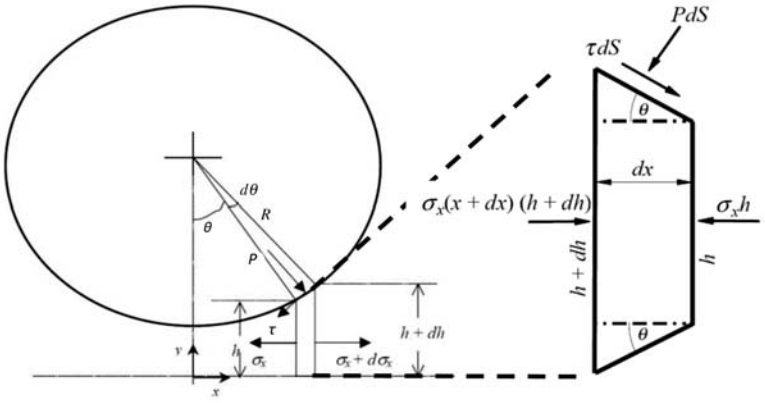


FIGURE 32.33 Force balance on a slab element during roller compaction and free-body diagram of slab. The stresses acting on the slab are described such that the roll pressure, P, and roll shear stress τ, act at the roll surface, which is defined by the angular increment, $d\theta$, and the slab surface in contact with the roll, dS. *Source: Adapted from Cunninghham JC. Experimental studies and modeling of the roller compaction of pharmaceutical powders. In:* Materials engineering. *Philadelphia, PA: Drexel University; 2005. p. 267. and Balicki M. Numerical methods for predicting roll press powder compaction parameters. Albi: Ecole Des mines D'Albi-Carmaux; 2003.*[109]

stress distributions and ribbon densities throughout the different positions in the roller compaction process. The primary assumptions applied in slab analysis are listed in Table 32.7.[35] Slab analysis has been commonly used in the metals industry to study rolling and extrusion for decades, making it an obvious avenue to pursue for roller compaction due to the similar nature of the processes.[88,108] In slab analysis, a plane stress and the existence of a slip region is assumed. However, the point at which different models diverge is the selection of the yield criterion, which defines the transition of elastic to plastic deformation in the material. It is also important to emphasize that models based on slab analysis require the nip angle as input, which may be experimentally determined or estimated using Johanson's model, and the powder density at the nip angle. Similar to Johanson's model, slab model results enable the prediction of ribbon densities for a set of input conditions (ie, material and process conditions), as well as provide a means to reverse-calculate necessary input conditions to achieve a target output (ie, ribbon density).

TABLE 32.7 Assumptions for Slab Method Models for Roller Compaction

Slab method assumptions
One-dimensional variation in field quantities (eg, stress and strain constant through slab thickness)
Deformation is considered plane strain (ie, strain in the roll width direction is zero)
Homogeneous deformation state exists in the slab element
Constant friction coefficient following Coulomb's friction law
Constant circular arc of contact
Negligible elastic deformation
Constitutive model is porous plasticity
Body forces can be ignored (eg, gravity)
Inertial effects are ignored

Katashinskii[93] proposed a yield strength, which was met by the difference of the principle stresses. Lee and Schwartz[96] introduced the Coulombian criterion and assumed that the yield stress and the cohesion in the Coulombian criterion were dependent on the state of material compaction, and thus, a function of the rolling angle. The slip region of the strip element was proposed to impart tangential forces to both rollers, which were set to be equal to the coefficient of the external friction multiplied by the normal force acting on the strip, enabling completion of the force balance. Katashinskii and Shtern[94,95] modified Katashinskii's[93] original approach to develop an equilibrium equation based on roll friction, normal stress acting in the rolling direction, and the roll pressure to describe the stress–strain behavior of a slab element experiencing shear deformation and densification. The continuity equation was used to account for powder compressibility. Roll friction was based on Coulomb's law and allowed to vary with rolling angle and assumed to be zero at the neutral angle; although it should be noted that there is no physical justification for such an assumption.[65] Input parameters such as initial feed stress acting in the rolling direction, rolling angle, neutral angle, coefficient of friction, initial relative density of the powder, gap and roll diameter are required from the user. Musikuin[97] developed a model for roller compaction that mirrored an approach commonly used for cold rolling of continuous metals, in which particle deformation was assumed to occur only in the nip

region. It was assumed that material "slips" at all angular positions except the neutral angle, θ_0, and once it entered the nip region, the particulate material behaved as a porous solid. The roll shear stress was defined by the limiting (slipping) condition and adhered to Coulomb's law. Musikuin then employed a simple, constitutive model for the material based on the difference between maximum and minimum stresses, which was the primary modification to similar models used for continuous metals. The principal stress was assumed to be zero at the bite angle and at the centerline of the rolls. Input parameters supplied to the model include the tapped density, which was the assumed material density at the bite angle, and the material density at the centerline of the rolls (ie, ribbon density). The shape of the roll pressure profile resulted in a combination of monotonic concave curves that met at sharp points at the neutral position. In order for the model to match maximum roll pressure data reported in literature, additional manipulations to model parameters were necessary. Thus, the primary limitations to Musikuin's model were the neglect of tangential stresses, assumptions that slippage only occurred at the roll surface, leading to a pressure profile shape that did not match that of experimentally measured profiles, and its reliance on manipulation of multiple model parameters to match experimental data. Dec et al.[65,99,100] introduced a slab model for roller compaction that utilized a yield criterion for metal powders that was originally proposed by Kuhn and Downey,[4] which is based on the relationship between Poisson's ratio and density. The nip angle value and the position of the neutral plane were determined from experimental data. The density profile for each compaction step was determined from compression test data for a corresponding mean stress. The predictive capabilities of this model were severely limited due to its heavy reliance on experimentally measured values as model inputs and the forced matching with the output density. More recently, Schonert et al.[43] applied the slab method to roller compaction by utilizing a normal-to-shear stress relation (ie, transmission coefficient) in a powder as a substitute for a yield criterion by representing the relative amount of pressure transmitted from σ_x to σ_y.

Shima and Yamada pursued an alternate route that incorporated principles applied in the mechanics of powder compaction and utilized an energy-based continuum model, in which the upper bound theorem was used to analyze the roller compaction of metal powders. The upper bound theorem is based on determining a probable velocity field for the material that minimizes energy dissipation attributed to both powder densification and friction at the roll/powder interface. The authors used a yield criterion that was originally introduced by Shima and Oyane,[87,98] in which a constant friction law dictated that frictional shear stress was constant and independent of normal stress. Initial assumptions for the values of gripping angle and velocity field were needed for the model input and then these values were refined through an iterative process, in which the rate of energy dissipation was minimized and the continuity equation was satisfied. The resultant density and strain rate distributions for the slab could then be calculated and used to determine stress distributions. Primary findings from this model were: (1) powder/roll frictional conditions had minimal impact on density distributions and, (2) gripping angle values significantly affected density distributions. Within this work, the authors presented seven case studies and identified rather significant discrepancies in several of the cases, with the specific causes for these discrepancies not discussed.

Numerous approaches have been used to apply slab analysis to the roller compaction operation with varying degrees of success, while demonstrating advantages and disadvantages of each approach. However, a primary drawback for all of the models built upon slab analysis is their dependence on experimental inputs, including the positions of the nip and neutral angles, as well as initial feed stresses, and thus limiting their predictive capabilities. As mentioned previously, these values are difficult to obtain experimentally, requiring specialized instrumented equipment. Such obstacles produce an additional limitation to the widespread use of the slab method.

32.3.3.1 Slab model comparisons with experimental data (instrumented roll)

Patel et al.[49] predicted MCC pressure distributions in a roller compaction process using a slab model, where nip angle was estimated using Johanson's theory. The model, which used input parameters generated from uniaxial compaction data, underestimated compaction pressure values by a factor of 1.5–2.5, although the predicted trends matched experimental data (Fig. 32.34). The discrepancy between the predicted and experimental results was attributed to (1) using an instrumented roll with a single load cell in the center of the roll, which did not account for pressure variation across the roll width, and it has been shown that roller compaction of MCC tends to produce nonhomogeneous pressure distribution profiles where pressures at the center of the ribbon are higher than those at the edges.[35,41,64,102] Therefore experimental pressure readings taken from the center of the roll would be expected to be higher than that predicted by a model that assumes uniform distribution of stresses across the ribbon width; (2) Nonuniform flow of powder in the nip region and through the rolls, yielding shear forces that do not occur during uniaxial compression; (3) The neutral angle, and maximum pressure, was assumed to be coincident with

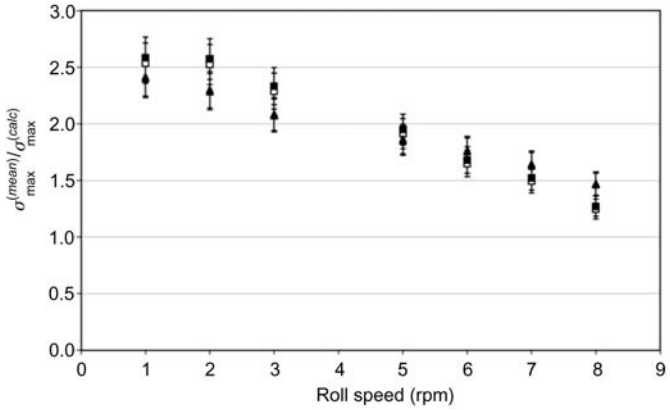


FIGURE 32.34 Ratio of the maximum measured pressure to predicted pressures, as a function of roll speed at roll gaps of □ 0.9 mm and ■ 1.6 mm.[49] *Source: ©2010, with permission from Elsevier.*

the minimum roll gap in Johanson's model, however, maximum pressure is known to occur before the centerline between the rolls (i.e., the minimum roll gap).[35] Overall, Patel et al.[49] found uniaxial compaction to provide a suitable first-order representation for roller compaction operations, although the authors warn that the model cannot accurately simulate nonhomogenous ribbon properties and shear effects that are present during real-life operation.

32.3.4 Finite element method

FEM is a numerical technique used to approximate the solutions of complex equations through the subdivision of a whole problem domain into simpler parts, or finite elements. Solutions are achieved by minimizing associated error functions for the collection of elements. To better visualize the concept of FEM, one can see how the connection of many tiny lines can be used to approximate a circle. FEM incorporates methods for connecting many simple element equations over many small sub-domains, the finite elements, to approximate a more complex equation over a larger domain.[110] FEM is an ideal resource for problems with complicated geometries, loadings, and complex material properties and has been found to be well-suited for systems of dense solids, where the material behaves in an elastic or elasto-plastic manner.[111] With respect to the roller compaction process, the deforming powder region is divided into a number of continuously distributed elements, or meshes. The mechanisms driving motion and deformation (ie, strain) of each of the elements are computed numerically according to a user-defined stress-strain constitutive relationship. The use of FEM to model tablet compaction had been reported in several studies, and thus many of the concepts that were applied in these previous works were translated to

TABLE 32.8 Assumptions for FEM-Based Models for Roller Compaction

FEM roller compaction model assumptions
Effect of interstitial air is ignored
Roll deformation due to roll pressure is ignored
Coulomb friction is the dominant frictional interaction at the powder/roll interface
Gravity and inertia are ignored
Screw feeder effect is approximated by stress at the entry angle

modeling of the roller compaction process.[18,112–114] Typical assumptions of FEM-based models for roller compaction are listed in Table 32.8.

FEM-based roller compaction models two-dimensional (2D)[35,41,64,65,89,101] and three-dimensional (3D)[64,102,103]) have enabled further insight into the stress and displacement fields that are typically generated during the unit operation, compared with Johanson's- and slab-based models, because they can account for through-the-thickness variation, more accurate representation of material behavior (typically using the Drucker–Prager/cap (DP/C) model[18,64,112–114]), frictional considerations, and feeder system effects. While friction for the roll/material is typically assumed to follow Coulomb's friction law with a constant friction coefficient, specific frictional considerations (ie, slip versus stick) can be directly evaluated in order to satisfy both equilibrium and compatibility and are not restricted by an upper bound argument. In addition, the effect of the feed system can be studied by varying the feed stress (constant[41,64,65,89] and oscillatory[35,64]) applied to the mesh in the rolling direction at the in-flow boundary. Early simulation attempts observed severe mesh distortions when implicit Lagrangian simulations were utilized.[35] Thus, the Arbitrary Lagrangian-Eulerian (ALE) method has since been used for its adaptive meshing capabilities. As an example, a schematic of a mesh designed for roller compaction analysis is shown in Fig. 32.35.[64] The mass and densities of the roll and material mesh were optimized to minimize both the inertial effects for the quasi-static deformation problem and computational time.

FEM simulations have been found to accurately reflect trends that had been previously observed experimentally, and thus confirm well-known empirical correlations. Typical simulation outputs include roll pressure, roll torque, roll shear stress, nip and neutral angles, and densification values. As expected, maximum roll pressures and exit densities were found to be significantly impacted by both feed stress and frictional coefficient values. Roll force, roll torque, and consequently, exit densities, all increased with increasing feed stress and frictional coefficients. However, in addition to verifying

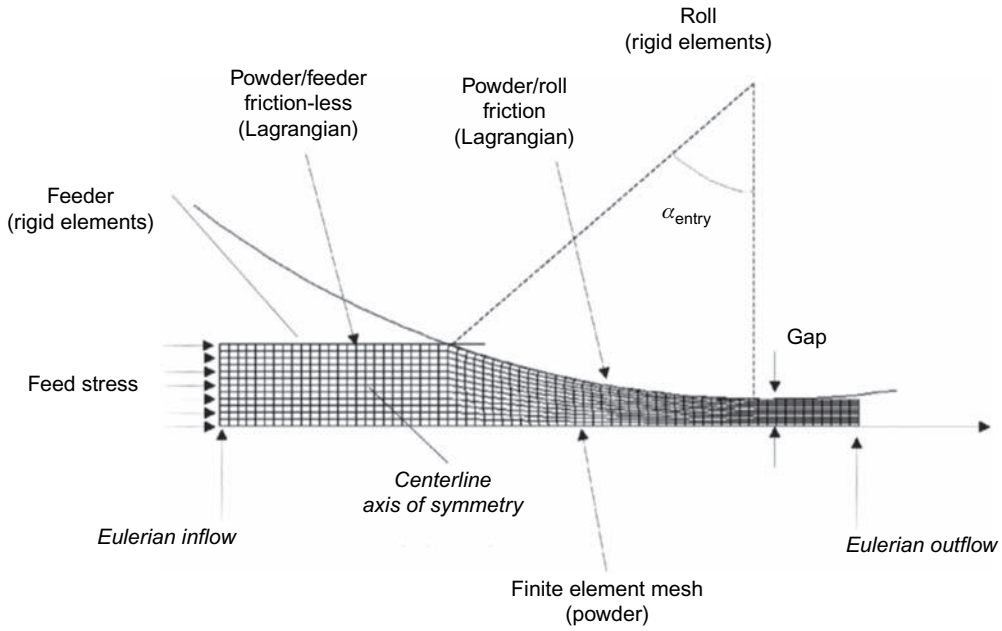


FIGURE 32.35 Schematic of finite element mesh for roller compaction with notation of axis of symmetry, Eulerian in-flow and out-flow, Lagrangian roll and feeder boundaries, and rigid bodies for the roll and feeder.[64] *Source: ©2010, with permission from Elsevier.*

known empirical correlations, FEM models have elucidated higher level principles regarding roller compaction mechanisms. For instance, the existence of slip and nip regions in the compaction zone was verified, as delineated by a marked increase in roll pressure and shear for rolling angles near the minimum roll separation. In addition, maximum pressures were predicted to occur 0.5–1.1 degree before the centerline between the rolls (0 degree),[64,65] in line with experimental observations.[35] FEM simulations also highlighted the probability of nonuniform velocity fields across the ribbon width, a result that the 1D-Johanson model cannot capture.[64,102] These nonuniform velocity fields were correlated back to nonconstant feed conditions, as well as potential frictional effects. To further probe this finding, the application of oscillating feed stress conditions (a common occurrence in screw feeders) has been shown to produce periodic variations in roll pressures and exit densities,[64] in line with experimental observations.[14,41] Simulations have also indicated that powders possessing larger friction coefficients, for a given feed stress, yield higher maximum roll pressures and maximum relative densities. Such findings can be physically explained because powders possessing higher friction coefficients grip the powder sooner in the roller compaction operation and thus undergo densification sooner, although there appears to be a limit to the impact of the friction coefficient on roll pressure, above which its effect becomes limited.[64]

To better understand the role of frictional effects on roller compacted products, Cunningham et al.[64] employed a 3D model to account for the presence of side seals on roller compaction units. In roller compaction operations, side seals are used to minimize leaking of powder to the roll edges in an attempt to reduce the amount of ungranulated material (ie, fines) in the final granulation. The side seals were found to induce frictional effects that resulted in lower roll pressure and densities near the edges, compared with the ribbon center. Material selection that increased side seal friction resulted in higher density variations throughout the transverse direction of the compact, in agreement with experimental observations. The model was also used to assess the impact of variable in-flow of material (another common experimental observation for screw feeders) along the roll width on roll pressure. Simulation results showed that both the side seal friction and nonuniform feeding to the rolls contribute to nonuniformity in roll pressure and density across the roll width and provided rationale behind experimental observations.

While qualitative comparisons have been made between FEM simulation results and experimental observations, few studies have quantitatively assessed the accuracy of FEM predictions. Cunningham[35] reported good correlations between predicted and experimentally determined nip angles for an MCC system. However, predicted ribbon densities, as a function of roll pressure, were found to overestimate experimental values by 15–20%. These discrepancies were attributed primarily to post-gap expansion of the MCC, which was not accurately captured by the simulation. Determination of typical expansion values for MCC and accounting for the

excessive expansion accordingly brought the predicted values much closer to the experimental values. Muliadi et al. went on to quantitatively compare the results of a 3D FEM simulation with experimental data.[103] Roller compaction experiments were performed using an air-powered piston feeder, as opposed to a more traditional screw feeder, in an attempt to apply a uniform feed stress to the powder. MCC was also used as the representative compound for this study. Similar modeling conditions and parameters to those used by Cunningham et al.[64] and Michrafy et al.[102] were applied (ie, ALE scheme, modified DP/C model for the material stress-strain constitutive behavior, and Coulomb friction to approximate powder/roll interactions). In this case, however, the inlet stress location was determined by monitoring the point at which the feed piston stroke reached its maximum. The roll gap corresponding to the maximum piston stroke was established using a laser distance sensor. Thus, no parameter-fitting was used in this method, making this FEM-based model truly predictive. By meticulously ensuring that the input parameters to the simulation were as accurate as possible, excellent correlations between the simulation and experimental results were achieved. Fig. 32.36 demonstrates the accuracy of predicted ribbon densities, as a function of ribbon position, to experimentally measured values for multiple-feed stress conditions. These results are also in alignment with predicted trends reported by the previously discussed 3D FEM models, demonstrating that ribbon density is highest in the middle of the ribbon and lower at the edges.[64,102] However, it should be noted that the sensitivity of the FEM results to the friction coefficient was highlighted as one of the primary weaknesses of FEM models. FEM simulations rely heavily on the user's experimental input. Therefore, appropriate parameters for feeding conditions and friction models (because processing energy is transmitted throughout the roll-material contact) and powder properties are imperative if these simulations are to be truly predictive.[101] An additional drawback to the proliferation of these FEM-based models is the requirement of expensive computer resources and expertise. However, once expertise in this technique is established, FEM-based analysis has demonstrated a promising ability to provide mechanistic insight into rolling compaction, and as a result, enable improvements in process and equipment optimization that may outweigh the costs of initial investment.

32.3.4.1 Comparison of FEM to Johanson's theory

Given the widespread use of Johanson's theory to model the roller compaction process, Muliadi et al.[89] quantitatively compared the results obtained from Johanson's theory with those simulated by a 2D FEM approach. MCC was selected as the model compound. Predicted nip angle, maximum normal stress, and maximum relative density values were compared for the two models. The predicted impact of select input values, such as roll powder friction coefficient and inlet normal stress (feed stress), on the model output is shown in Fig. 32.37. Nip angle predictions by both approaches followed similar trends and were within ~25% of each other. However, the relative densities and maximum normal stresses predicted by Johanson's model were consistently higher than those predicted by FEM. Interestingly, Johanson's theory predicted relative density values greater than 1 for all of the conditions assessed in this study. These physically unattainable values were attributed to the fact that Johanson's theory, as a 1D model, was unable to capture the 2D velocity gradient experienced by the powder as it moved through the rolls. This 2D velocity field is also believed to enable some portions of the material to remain in a condition of "slip" even in the nip region. The inability

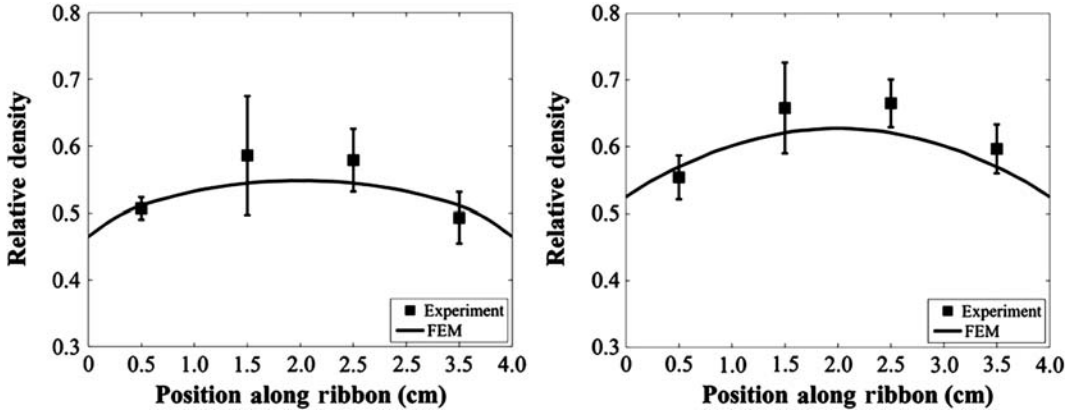


FIGURE 32.36 Comparison of FEM-predicted and experimentally measured ribbon density distributions for a feed stress of 132 kPa. Left and right plots correspond to a hydraulic roll force of 5.5 and 9.2 kN, respectively. A ribbon position of zero is farthest away from the compactor roll shaft support and a ribbon position of 4 cm is closest.[103] *Source: ©2013, with permission from Elsevier.*

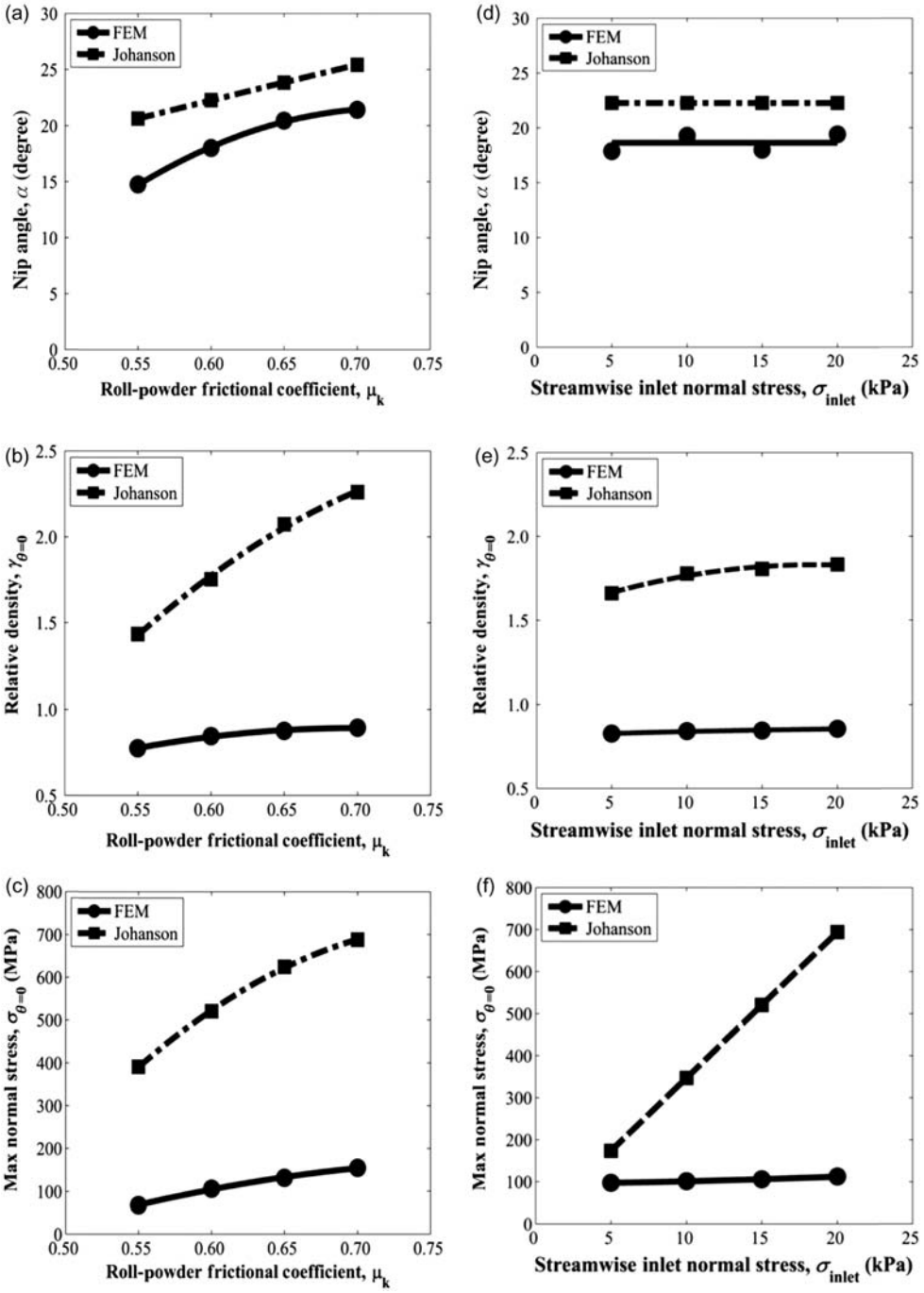


FIGURE 32.37 FEM-predicted (a) nip angle, (b) maximum relative density, and (c) maximum normal stress values as a function of roll-powder friction coefficient. FEM-predicted (d) nip angle, (e) maximum relative density, and (f) maximum normal stress values as a function of inlet normal stress. Additional parameters were σ_{inlet}, roll gap, roll diameter, and inlet height were 15 kPa, 2, 120, and 15 mm, respectively.[89] *Source: ©2012, with permission from Elsevier.*

of Johanson's theory to capture the 2D velocity gradient led to an overprediction of mass flow rate of material into the nip region. Such a phenomenon would be exaggerated for highly compactable materials, such as MCC. In terms of maximum normal stress, Johanson's theory predicted stresses larger than those by FEM. However, agreement between the two approaches improved for powders that were more compressible, possessed smaller effective friction angles, larger roll powder friction coefficients, and lower feed stresses.

DEM modeling of the roller compaction process will not be discussed in this chapter because the few

attempts that have been reported have not been as successful as FEM models.[111] Odagi et al.[91] modeled the roller compaction operation using DEM, but reported pressures in the nip region to be several orders of magnitude smaller than experimentally obtained values. In general, it appears that a continuum approach, as opposed to a discrete approach, may be better suited to roller compaction conditions, especially where large pressures are generated, such as in the nip region.

32.3.5 Neural networks and other artificial intelligence approaches

Pharmaceutical applications of neural networks, a form of artificial intelligence, have included direct compression of tablets and fluid bed granulation, in addition to roller compaction. These artificial intelligence approaches center on establishing cause and effect relationships within data sets and then using those relationships in "reverse" order to determine the process conditions necessary to achieve user-specified outputs. Such techniques are useful when gaps exist in experimental measurements of important process parameters or when complexities in both physical process geometry and mechanical behavior of powders impede process modeling efforts. However, it should be noted that these results are not founded upon physical or mechanistic principles of the system, but empirical correlations gleaned from the provided data sets. For this reason, the largest disadvantage of these models is their limited predictive capabilities to the training set conditions. Predictions for conditions outside of the training sets must be approached with caution.

For roller compaction systems, neural net models have been used to predict how process parameters[39] and formulation selection[104] can impact downstream operations. For instance, roller-compacted ribbons possessing higher densities tend to yield stronger, harder granules, with lower granule friability. While strong granules are desirable for their attractive flow properties, excessively strong granules may not readily disintegrate and may not compact well, adversely affecting dissolution and tablet strength, respectively. Optimization of formulation and process input variables, with respect to granule friability, tablet strength, and disintegration, is not trivial because these target values move in opposite directions of each other. Such complicated relationships would be difficult to balance using a mechanistic model, thus making these scenarios ideal for a neural network. Inghelbrecht et al.[39] used a neural net model to accurately predict granule friability and particle size of maize starch based on user-specified processing conditions (feed screw speed, roll pressure, and roll speed). Turkoglu et al.[104] utilized a neural network to guide formulation decisions for an acetaminophen tablet, for which binder grade and concentration, extra-granular material consideration, and number of roller compaction passes were optimized to yield granules of acceptable friability and tablets possessing adequate strength and disintegration properties. More recently, Mansa et al.[105] demonstrated the ability of a commercial software package that utilized neural network, genetic algorithms, and fuzzy logic to successfully predict ribbon densities from user-specified process conditions (roll speed, roll gap, and hydraulic roll pressure) and formulation considerations (material type, particle size distribution, powder densities, compressibility, tensile strength, friction angles, and flow function). The true value of the model, however, was demonstrated not in its ability to predict process outputs from user-specified inputs, but the ability of the model to outline the necessary inputs (process and formulation conditions) to achieved target outputs (ribbon density). Model-predicted conditions were successfully verified experimentally.

32.4 APPROACHES TO DEVELOPING A ROLLER COMPACTION PROCESS

A suggested approach to developing a robust roller compaction process, in accordance with QbD principles, is summarized in Fig. 32.38 and described in further detail below. The establishment of appropriate critical material attributes, CPPs, and quality attributes for both the ribbons and granules early on in the process enables an efficient and robust assessment of the operation and lends the process well to future technical transfer and scale-up activities.

32.4.1 Material assessment: appropriateness for roller compaction

When selecting roller compaction as the granulation approach for a project, consideration should first be given to the appropriateness of the formulation for manufacture using roller compaction. As previously described, blends with drug loadings <60% are amenable for roller compaction. Material characterization of the feed powder should be assessed in to provide insight into potential manufacturing challenges (Refer to Table 32.1). For instance, blends with low bulk density or poor flow properties may not efficiently pass through the rolls. Insufficient compactability may indicate a potential for poor tablet compactability of resultant granules. The propensity for a material to adhere to surfaces (ie, roll surface) is another important trait to consider. Material adhesion challenges, typically

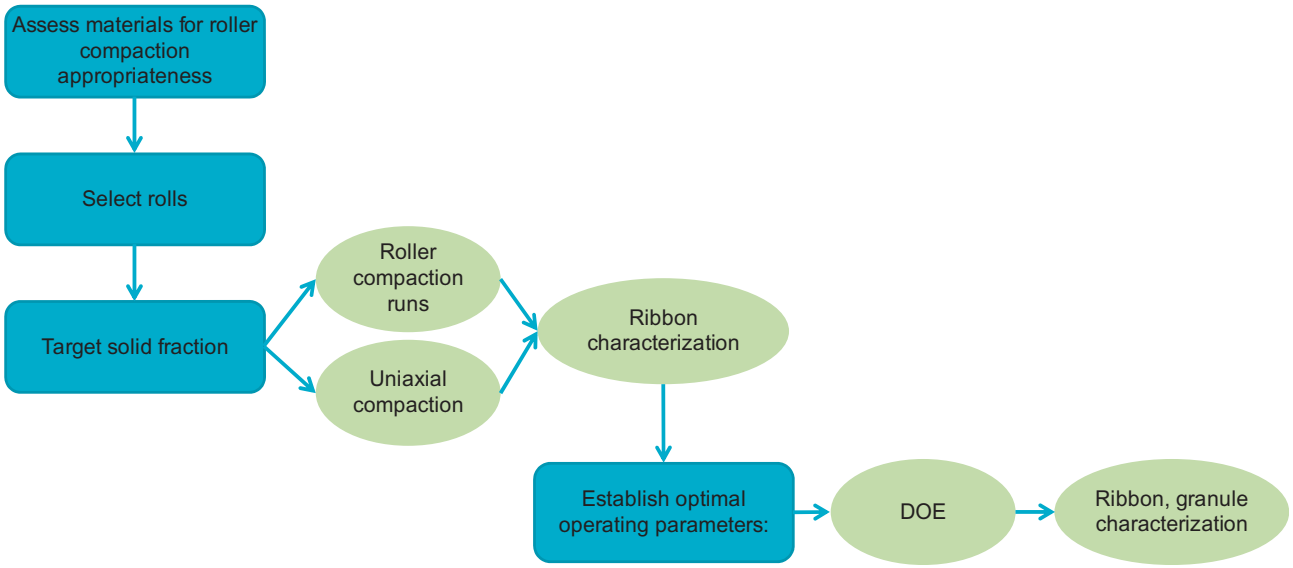


FIGURE 32.38 Flow diagram describing suggested approach to roller compaction process development.

associated with APIs rather than excipients, are difficult to assess because a robust characterization method has not been established. When material is observed to adhere to surfaces during early development work, proper excipient selection is often the first-line strategy to mitigate adhesion challenges, which typically includes modification of lubricant selection or level; however, in cases where acceptable mitigation is not achieved, changes to API form, particle size or particle morphology may be considered. If powder adhesion still cannot be properly rectified, alternative granulation techniques, such as wet granulation methods, are suggested.

32.4.1.1 Roll selection

As mentioned in the previous section, roll surface design is known to impact powder flow through the roller compactor. Patterned roll surfaces (ie, knurled, serrated) assist in "grabbing" powder and facilitating powder flow between the rolls. Roll surface design may also be a critical factor with powders that have a high tendency to stick or adhere to surfaces. The finish of the roll surface can affect roller compaction properties, such as the nip and entry angles during compaction, and thus the cumulative pressure applied to the powder. Therefore, it is advantageous to select the rolls at an early stage to provide consistency throughout the development process. Different roll designs and combinations (ie, smooth upper and lower rolls vs smooth upper/knurled lower roll) should be evaluated. When conducting preliminary experiments, typical roll gaps (2.0–3.0 mm) and 3–4 different roll pressure values should be evaluated. Feed screw performance can be a good indicator of powder flow performance; for example, if powder is not effectively drawn through the rolls and a back pressure is being created, then the load (ie, torque) on the feed screw may increase, and if using a closed feedback loop between the roll gap and the feed system (ie, gap control), then the feed screw speed will be significantly increased as the instrument attempts to feed more material between the rolls to open the roll gap.

32.4.1.2 Optimizing roller compaction parameters: targeting a ribbon solid fraction

The next objective is to identify the target range for ribbon solid fraction, which will relate to key roller compaction process parameters and the remaining compactability of the granule, which is important for subsequent tablet manufacture. It is recommended to produce ribbon or compacts at three or four different solid fraction values within the range 0.6 and 0.8 (refer to Fig. 32.6). The lower end of the range should be defined by the lowest solid fraction ribbon that can be milled to produce desirable granules, without a significant increase in fines content. Historically, a target ribbon thickness is selected (within the range 2.0–3.0 mm is typically acceptable) and the feed rate and roll pressure are manipulated to achieve acceptable solid fractions (0.6–0.8) while maintaining the target ribbon thickness. Sample analysis should include solid fraction and tensile strength measurements. Subsequently, the ribbons are milled to produce granules, which should

be assessed for compactability to determine suitability for tablet manufacture. Additional granule characterization includes particle size and powder flow analysis. Generally, ribbons of lower solid fraction result in granules comprising smaller particles that retain good compactability but may display poor powder flow; whereas higher solid fraction ribbons result in larger particles with relatively less compactability that are likely to possess desirable flow properties. Ribbons of suitable solid fraction produce granules that result in sufficiently strong tablets that can be manufactured on a high-speed tablet press while maintaining good tablet weight control. Ideally, a tablet of 0.85 solid fraction should have a tensile strength of greater than 1.7 MPa[115] and surpass tablet friability standards as defined by USP <1216>; and tablet weight should generally be maintained within 95–105% of the target weight throughout the manufacturing run.

More recently, small-scale compaction techniques have been used to predict target ribbon solid fractions before conducting any work on a roller compactor. When material or development time is scarce, the use of a uniaxial compaction method to simulate roller compaction (technique described in detail in Section 32.2) is a useful tool for producing compacts over a range of solid fractions. Simulated roller compaction ribbons may be prepared on a compaction simulator; these compacts may then be tested using characterization methods identical to those used for actual roller compaction ribbons. The simulated compacts may also be milled to yield granules, to emulate those prepared by roller compaction. The use of this small-scale simulation method provides >70% reduction in material resources compared with traditional development approaches.

32.4.1.3 Optimizing roller compaction parameters: establishing the design space

Once the target solid fraction has been identified, the acceptable solid fraction range should now be determined. The necessary operating parameters required to achieve the target solid fraction may be estimated using one of the roller compaction models described in the previous section. In lieu of modeling results, a design of experiment (DOE) should be conducted based upon knowledge from previous products. If modeling results are available, the DOE should be designed around the predicted edges of failure.

Key product attributes for the roller compaction ribbon are solid fraction, thickness, and tensile strength (see Section 32.1 for more detail). However, it should be noted that ribbon tensile strength is not commonly monitored during the commercial manufacturing process: ribbon solid fraction and thickness, which relate to tensile strength, are more easily measured attributes and thus more amenable to commercial manufacturing. Granule properties (eg, particle size distribution, flow) may also be measured offline to provide an indication of the roller compaction process. Key process parameters that impact ribbon and granule quality are roll pressure, feed pressure, and roll gap. Feed pressure is primarily dictated by feed screw speed and is typically controlled via an internal feedback loop rather than being adjusted by the operator. As such, the DOE is often built around targeted roll pressure and roll gap values and the feed pressure is allowed to float. A two-factor, three-level DOE (2^3) investigating roll pressure and roll gap is generally considered to be sufficient, especially as a first-pass assessment of the operation. During this DOE, throughput requirements (at commercial manufacturing levels) should be considered when selecting roll speed values. The parameters from the initial study, which was used to identify a suitable target solid fraction, may be used as the center points, plus upper and lower range values. Upper and lower roll pressure range limits might be based on those from the initial study, or an arbitrary range can be used, such as ±10 bar from center. The operating range for roll gap should be at least ±0.2 mm from target; thus a minimum operating range would span 0.4 mm. Ribbon samples should be analyzed to determine solid fraction and, if capabilities allow, tensile strength.

32.4.1.4 Output characterization

Ribbon quality is considered to be an indicator of granule quality. However, it is ultimately granule quality that is most critical for downstream operations. Thus, ribbon samples should be milled to produce granules and subsequently assessed for compactability and flow. Particle size distribution data is also valuable, and can typically be related to granule flow performance. When milling ribbons, especially throughout a DOE investigation, it is recommended to use a single mill screen aperture size to avoid confounding effects from changes to process parameters. Recommended screen mesh sizes are in the range of 0.8–1.2 mm to break up loose aggregates without reducing the granules back to fines. It is also recommended to test the impact of different screen sizes on granule quality to demonstrate process robustness once the working ranges for roll pressure and gap have been determined. Process robustness against milling conditions has late-stage development and scale-up implications because many commercial roller compactors are integrated with a mill system, and thus milling conditions may not be exactly translated.

32.5 ILLUSTRATIVE EXAMPLE DETAILING THE TYPICAL DRUG PRODUCT DEVELOPMENT PROCESS FOR A ROLLER COMPACTED PRODUCT

Formulation A, comprising 35% active drug (Tables 32.9), required the development of a dry granulation method using an Alexanderwerk WP120 roller compactor. Selection of target processing parameters for the manufacture is detailed below. While not detailed in this example, similar studies may be executed to understand the impact of API and excipient material attributes, in addition to process parameters, on product performance.[8]

32.5.1 Selection of rolls

32.5.1.1 *Experimental set-up*

Smooth/Smooth and Smooth/Knurled upper/lower roll configurations were evaluated to assess the impact on roller compaction performance. The operating conditions for roller compaction manufacture are detailed in Table 32.10.

32.5.1.2 *Outcome*

When manufacturing using two smooth rolls and gap-controlled operations, feed screw speeds and load readings reached 100% capacity; whereas the smooth/knurled roll configuration resulted in acceptable and stable feed screw loads of approximately 60% capacity, while maintaining a consistent roll gap. Therefore, the smooth/knurled roll configuration was selected because it better facilitated good powder flow through the roller compactor compared with the smooth/smooth roll configuration.

TABLE 32.9 Composition of Formulation A

Formulation composition		
	Role	Wt %
Intra-granular		
Active drug	API	35.0
Lactose monohydrate	Filler	30.0
Microcrystalline cellulose	Filler	30.0
Silicon dioxide	Glidant	1.0
Croscarmellose sodium	Disintegrant	2.0
Magnesium stearate	Lubricant	0.5
Extra-granular		Wt %
Croscarmellose sodium	Disintegrant	1.0
Magnesium stearate	Lubricant	0.5

TABLE 32.10 Roller Compactor Operating Conditions

Property	Target values
Hydraulic roll pressure (bar)	35, 70
Roll gap (mm)	2.2
Operating method	Gap-control

32.5.2 Selection of operating parameters

32.5.2.1 *Experimental set-up*

- Small (~1–2 kg) batches were manufactured on the roller compactor to manufacture ribbons at varying solid fractions (0.65, 0.70, 0.75, and 0.80) with a target ribbon thickness of 2.2 mm. Subsequent ribbons were milled using the in-line mills of the Alexanderwerk roller compactor. Ribbon manufacture could also be executed using a compaction simulator, in the event that material is limited.
- Granules were compacted into tablets using a compaction simulator
- Two-factor, three-level DOE was conducted to optimize acceptable roll pressure and roll gap ranges (Table 32.11).

32.5.2.2 *Outcome*

- Roller-compacted ribbons of 0.65 solid fraction were weak, possessing tensile strengths of 0.7 MPa, and resultant granules of the 0.65 solid fraction ribbons possessed a higher fines content compared with the other lots. Ribbons of 0.70 solid fraction possessed tensile strength of 1.8 MPa, exceeding the recommended target of >1.0 MPa at 0.7 solid fraction (Figs. 32.39 and 32.40).
- Tablets manufactured from 0.80 solid fraction ribbons were weaker than those manufactured from the other three lots of granules (Fig. 32.41). Granules from 0.65, 0.70, and 0.75 solid fraction ribbons displayed satisfactory compactability; however, the higher solid fraction granules displayed better mass flow in the tablet press hopper, which is attributed

TABLE 32.11 Experimental Plan for Two-Factor, Three-Level DOE to Investigate Roll Pressure and Roll Gap Range

Run	Hydraulic roll pressure level	Roll gap level	Hydraulic roll pressure (bar)	Roll gap (mm)
1	High	Low	74	2.0
2	Low	High	54	2.4
3	Center	Center	64	2.2
4	Low	Low	54	2.0
5	High	High	74	2.4

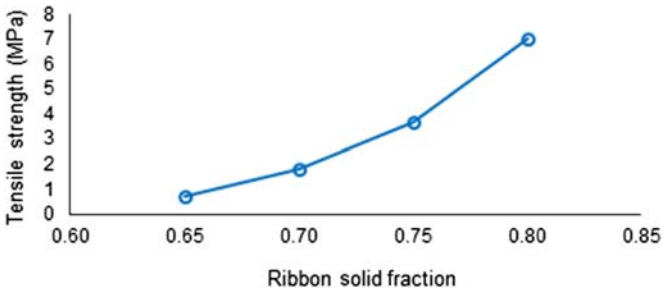


FIGURE 32.39 Assessment of ribbon tensile strength over the range of 0.65–0.80 solid fraction. Analysis performed using a three-point bending flexural test.

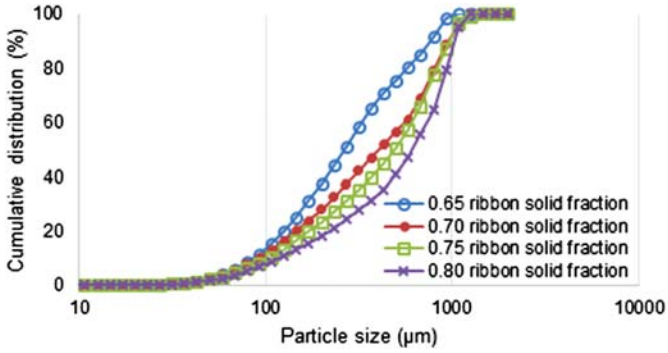


FIGURE 32.40 Cumulative particle size distribution for granules from ribbon of 0.65, 0.70, 0.75, and 0.80 solid fraction. Analysis performed using a Sympatec GmbH QicPic Particle Size Analyzer.

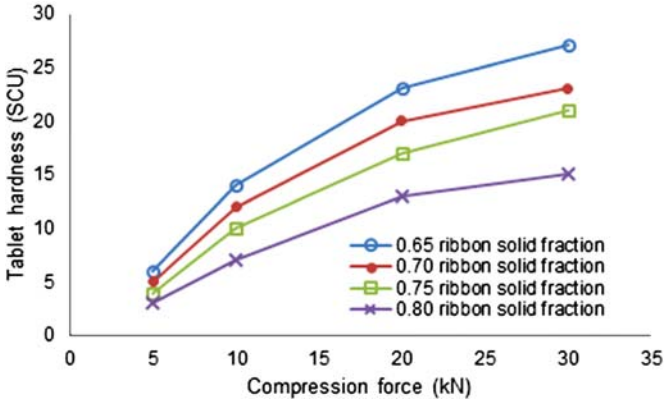


FIGURE 32.41 Compaction profiles generated using granules from ribbons of 0.65, 0.70, 0.75, and 0.80 solid fraction. Tablets manufactured using a Medel'Pharm Stylcam 100R press replicator with 9.00 mm round, biconcave tooling.

to differences in particle size distribution. During this evaluation, a hydraulic roll pressure of 64 bar was used to produce ribbon of approximately 0.75 solid fraction, therefore this roll pressure was set as the target for further work.

- A DOE was executed, in which a hydraulic roll pressure of 64 bar and roll gap of 2.2 mm were selected as center points and expected to yield ribbons of ~0.75 solid fraction based on data generated in the previous stage of work. The operating space was investigated by testing above and below the center points for roll pressure and roll gap, investigating high/high, low/low, high/low, and low/high combinations.
- Experimental results are listed in Table 32.12. Highlights from the results include:
  - Ribbons manufactured using center-point parameters averaged 0.74 solid fraction, comparable with the target of 0.75 solid fraction. The combination of low roll pressure/high roll gap resulted in the lowest solid fraction of 0.71 and the combination of high roll pressure/low roll gap resulted in the highest solid fraction of 0.77.
  - Based on granule flow characterization, ribbons of higher solid fraction produced better-flowing granules.
  - Higher solid-fraction ribbons produced less compactable granules, which resulted in tablets of lower hardness compared with those from ribbons of lower solid fraction.
  - All tablets exhibited less than 1% weight loss after 100 revolutions in the friabulator, which is acceptable.
  - All tablets achieved the dissolution target of Q80 release within 30 minutes.

Final conclusions: Small-scale batches enabled the establishment of:

- Target quality attribute for the ribbon: solid fraction of 0.75; and
- Target process parameters were set to the following: target hydraulic roll pressure was set at 64 bar with an operating range of 54–74 bar; and the target roll gap was 2.2 mm with an operating range of 2.0–2.4 mm.

32.6 SCALE-UP CONSIDERATIONS OF ROLLER COMPACTION

Pharmaceutical scale-up is commonly thought of as the process by which batch size is increased. This can be accomplished by enlarging the physical dimensions from lab- to pilot- to plant-scale or by increasing the output from a certain piece of equipment.[116] Roller compaction is a unit operation that readily lends[6,7] itself to scale-up by either method. Through the use of continuous processing, larger amounts of powders can be compacted using the same piece of equipment used

TABLE 32.12 Results of Two-Factor, Three-Level DOE Investigating Roll Pressure and Roll Gap Range; Solid Fraction Was Measured Using a Micromeritics GeoPyc 1360 Pycnometer; Granule Flow Was Measured Using an Erweka Powder Flow Tester; Tablet Friability Was Assessed at 100 Revolutions Using Apparatus Compliant With USP <1216>; and Dissolution Testing Was Performed Using USP Type 2 Apparatus at 50 rpm.

Run	Roll pressure/gap (center/low/high)	Ribbon solid fraction	Granule flow (g/s)	Tablet hardness (SCU)	Tablet friability (% weight loss)	Tablet dissolution
1	H/L	0.77	4.5	17	<1	Q80 @ 30 min
2	L/H	0.71	3.2	22	<1	Q80 @ 30 min
3	C/C	0.74	3.8	20	<1	Q80 @ 30 min
4	L/L	0.73	3.6	21	<1	Q80 @ 30 min
5	H/H	0.76	4.2	18	<1	Q80 @ 30 min

for smaller scale batches by running for a longer period of time. The two main advantages of continuous processes are that ease of scale-up for larger batches and a 24-hour automatic production line is possible.[117] For example, a roller compaction process could be scaled-up from a 40 kg batch to a 400 kg batch by simply increasing the batch run time from 1 to 10 hours.

Ideally, when scaling up by enlarging the physical dimensions of the roller compactor from one production scale to another, the equipment should be similar geometrically, dynamically, and kinematically.[117] The geometric condition is fulfilled when the ratio of physical dimensions between the small-scale and the scaled-up version are constant. Dynamic similarity is seen when the ratio of forces exerted between matching points in the two roller compactors are equal. Finally, kinematic similarity is met when the ratio of velocities between matching points in both systems are equal.[117] In reality, the scale-up process is more complicated because the equipment ratios between different scales may not match exactly. During scale-up consideration, regulatory guidance on manufacturing changes/similarity must be taken into account. Scale-up to larger roller compactor equipment is typically considered a "level 1 equipment change," according to the Food and Drug Administration's (FDA) Scale-Up and Post Approval Changes guidance document for immediate-release solid oral dosage forms (SUPAC-IR)[118–120] because roller compactors operate on the same operating principles and have the same design. Batch size increases must also be considered, since a level 1 change occurs when the production batch is up to 10-times larger than the pilot batch size. A level 2 batch size change occurs when the scaled-up batch size is greater than 10-fold for equipment operating on the same operating principles and design.[119,120]

Apart from the regulatory requirements, it is important to understand the effects of scale-up on the compacted ribbon and subsequently, the final dosage form. As discussed in previous sections of this chapter, the material attributes of interest from a roller compaction process are typically ribbon solid fraction, ribbon tensile strength, granule particle size distribution, and granule bulk density. During scale-up, it is generally accepted that if ribbon thickness and ribbon solid fraction are translated across scales, then granule attributes, and thus performance, will also be comparable. Ribbon and granule attributes are generally evaluated offline for confirmation of scale-up success. Ribbon density can be determined offline using an arc punch or GeoPyc or online using a PAT application, such as NIR spectroscopy as discussed in previous sections. Granule properties, such as particle size distribution, can be measured offline using a sieve method, light scattering technique (static or dynamic light scattering), or image analysis technique (QICPIC, Morphologi G3). Online particle size options include focused beam reflectance measurements or optical systems.[121,122] Scale-up focuses on the process parameters of roll pressure, roll gap, and roll speed because these parameters can have an effect on ribbon and granule attributes. The mill speed should also be considered as this can influence the milled particle size distribution. Once the development design space has been mapped to the commercial scale design space, the need to continue to measure these product attributes will be dependent upon the process control strategy.

The large-scale design space may be established by a few different methods, as listed in Table 32.13. The first option is to use historical experience to identify process parameters that are likely to produce ribbons and granules of acceptable quality. The issue with this approach is that it requires large amounts of material to perform a full DOE to translate the small-scale design space to the large scale. The rapid pace of pharmaceutical development and the rising costs of new active ingredients make this approach unattractive for most product scale-ups. The second approach is to use a surrogate formulation, typically a placebo formulation, to execute an experimental design to map the design space of the

TABLE 32.13 Scale-Up Approaches

	Historical experience	Surrogate mapping	Modeling

How it works	Trial and error approach, using historically known conditions as a starting point	• Execute DOE at commercial scale using surrogate material, often placebo
		• Apply DOE results to active blend
			• Execute experiments to collect model input
			  ◦ Limited roller compaction runs at small-scale
			  ◦ Material characterization
			• Using modeling results to predict roller compaction design space
			• Spot-check predictions at large-scale to verify modeling predictions

Considerations	Requires large amounts of material, as an entire DOE executed commercial scale	Results most accurate for low drug load formulations and when placebo blend is highly similar to active blend	• Requires extra effort initially, compared with previous scale-up methods, but overall uses significantly less material, resource • Ensure modeling assumptions are well understood

small- and large-scale roller compactors. The use of a common placebo formulation is an acceptable practice because the information gathered can be extrapolated to an active containing formulation. This approach saves the cost of running large-scale batches with active ingredient and, when paired with historical knowledge from previous successful scale-ups, can be used to define the large-scale design space. When considering the placebo formulation it is important to remember that the closer the placebo is to the real product's formulation the more accurate the scaled-up parameters will be for the commercial process. It is recommended to spot check the design space at center point and the extremes to ensure ribbon and granule properties align with product properties produced by the small-scale design space.

The third approach utilizes modeling techniques outlined in the previous section of this chapter. In this approach, a material characterization measurements are required up front to supply the inputs for the model and a limited number of experimental runs on small-scale equipment are required to validate the model. One of the key model inputs, which should be understood on both scales, is the relationship between hydraulic pressure and applied roll force. This relationship can be measured by the use of an instrumented roll or force cradle. Once the limitations/assumptions of the model are understood, it can be modified to generate a design space for the large scale. It is recommended to spot check the design space predicted by the model at center point and the extremes to ensure ribbon and granule properties align with product properties produced by the small-scale design space. This approach requires more up-front work, but requires less material and resources to be successful. In addition, the model can then be confidently applied to subsequent product scale-up activities.

32.7 ILLUSTRATIVE EXAMPLE DETAILING A POSSIBLE APPROACH TO SCALING-UP A ROLLER COMPACTION PROCESS

A combination of scale-up approaches was used to determine appropriate roller compaction conditions when scaling up from a roller compactor fitted with 120-mm-diameter rolls to a roller compactor fitted with 200-mm-diameter rolls.

32.7.1 Experimental design

1. Selection of material to be used for scale-up study
   a. To minimize API costs, a surrogate roller compaction placebo formulation was used to mimic a formulation with 4% wt drug load.
   b. Composition of the placebo was 1/1 MCC/lactose (anhydrous) and contained relevant amounts of disintegrant, glidant, and lubricant
2. Selection of target ribbon solid fraction
   a. A target ribbon solid fraction of 65–77% was selected, based on results obtained from a small-scale roller compaction simulation (manufacture of compacts using a uniaxial compaction simulator).
   b. The selected ribbon solid fraction range is in line with previously documented experience; thus, resultant granules were expected to possess sufficient compactablility for tabletting.
3. Selection of design space
   a. 12-run DOE devised to characterize the design space for both roller compactors (Table 32.14). The experiments were designed such that three separate analyses could be performed:
      i. Two-factorial design (runs 1–8) used to establish the impact of roll gap and roll

TABLE 32.14 Design of Experiments for Roller Compaction Runs

Run	Roll speed (mm)	Hydraulic roll pressure (bar)	Roll gap (mm)	Mill speed (rpm)
1	Center	Center	Center	Center
2	Center	Low	Low	Center
3	Center	Low	High	Center
4	Center	High	Low	Center
5	Center	High	High	Center
6	Center	Center	Low	Center
7	Center	Center	High	Center
8	Center	Center	Center	Center
9	Low	Center	Center	Center
10	High	Center	Center	Center
11	Center	Center	Center	Low
12	Center	Center	Center	High

pressure, which is typically found to be main factors to effecting ribbon density or tensile strength.

ii. Runs 1, 8, 9, and 10 to assess the impact of roll speed on ribbon density and granule properties.

iii. Runs 1, 8, 11, and 12 to assess the impact of mill speed on granule properties.

iv. Runs 1 and 8 at center point for all parameters to assess the reproducibility of the impact of process parameters on ribbon and granule properties.

b. Roll speed range determined by commercial throughput considerations

c. Milling parameters selected based upon historical experience

d. Ribbon density and granule particle size distribution were selected as the material attributes of interest.

32.7.2 Results

The DOE was executed using the surrogate placebo blend on both the pilot scale (120-mm rolls) and commercial scale (200-mm rolls) roller compactor.

- Ribbon density data showed that the commercial roller compactor required higher hydraulic pressure set points to generate a ribbon with a comparable density with that of the pilot-scale unit. This was expected since the rolls of the commercial unit were wider, and thus, a higher force was required to maintain the same compressive force per cm of roll width. The design space for ribbon density is given in Fig. 32.42.
- Particle size distributions for the two roller compactors were equivalent, as seen in Fig. 32.43. The data from all 12 conditions were averaged together and plotted along with the minimum and maximum observed results. As shown, the distributions completely overlap. In addition, both roller compactors exhibited the same relationships for changes in roll speed and mill speed. Statistical analysis of the results showed that neither roller compactor exhibited a relationship for roll speed on ribbon density or had a significant relationship for particle size with regards to changes in mill speed.
- Granule bulk and tapped density values were similar for all DOE conditions.

Overall, this scale-up approach was shown to be successful for providing suitable process parameters to

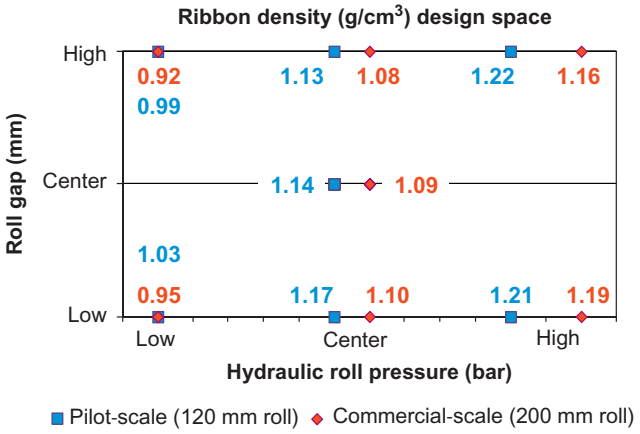


FIGURE 32.42 Ribbon density design space map.

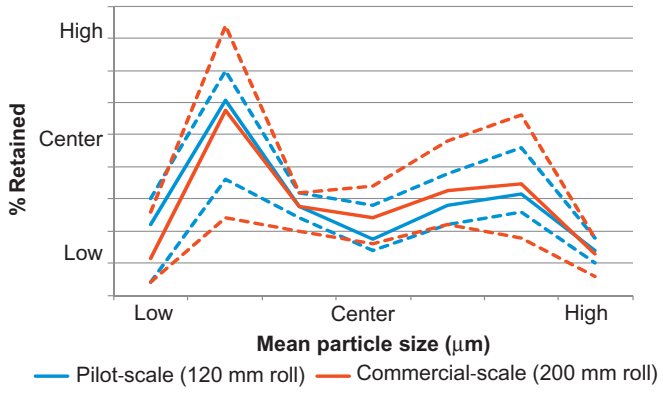


FIGURE 32.43 Granule particle size distribution. Dashed lines indicate minimum and maximum observed results and solid line is average.

TABLE 32.15 Common Roller Compaction Challenges and Mitigation Approaches

Problem	Possible cause	Remedies
Sticking to compaction rolls	Insufficient lubricant level	Add lubricant; typically 0.2–0.5%; external lubrication a possibility
	Excessive moisture	Reduce moisture content; Dehumidify environment
	Poor scraper adjustment	Decrease clearance between rolls and scrapers; Replace worn scrapers
	Excessive force	Decrease roll force
	Melting of product	Cool rolls to decrease temperature; use water, nitrogen
	Improper roll configuration	Use smooth or grooved rolls
Soft compact/friable granulation	Improper formulation	Increase compressible binder; Decrease lubricant level
	Compaction force too low	Increase compaction force
	Compaction (dwell) time too short	Decrease roll speed
	Excess blend time for powder and lubricant	Decrease blend time
	Raw material variation	Monitor key raw material specifications
Excessive fines	Air entrapment	Use deaeration system
	Leakage of powder	Tighten seals (top/side); replace worn seals
	Poor powder compressibility	Increase compressible binder; screen/sieve product to desired particle size
	Improper granulation/mill screen	Use coarser screen for milling compacts
	Improper granulation/mill speed	Decrease mill speed
	Low compaction force	Increase roll force
Low bulk/tap density	Compaction force too low	Increase roll force
	Particle size too large	Use finer screen/plate to mill the compacts
	Challenging material	May need to pass material through roller compactor several times
Poor flow characteristics	High fines content	See remedies for "excessive fines"
	Particle size too small	Use coarser screen/plate to mill the compacts
Poor granule dissolution	Compaction force too high	Decrease roll force
	Improper formulation	Add a disintegrant (ie, starch, MCC)
Drug release profile—too slow	Poor formulation-slow release	Increase %hydrophilic polymer; use polymer with lower MW; Decrease %lubricant
Drug release profile—too fast	Poor formulation-fast release	Increase hydrophilic polymer (ie, HPMC); use polymer with higher MW
Low throughput/production rate	Feed rate too slow	Increase screw/roll speed ratio
	Improper screw design	Use alternate screw (double flight)
	Air entrapment	Use a deaeration system
Poor product stability/discoloration	Heat generation during compaction or milling	Decrease compaction force; use water-cooled rolls; use nitrogen cooling
Improper ribbon thickness	Poor scalability of the screw/roll speed ratio	Increase screw/roll speed ratio to increase thickness; decrease screw/roll speed ratio to decrease thickness
	Improper roll gap setting	Adjust roll gap to desired thickness

(*Continued*)

TABLE 32.15 (Continued)

Problem	Possible cause	Remedies
Compaction in hopper/top seal	Improper screw/roll speed ratio	Increase roll speed; decrease screw speed
	Improper screw/hopper clearance	Increase clearance between the screw feeder and side walls of the hopper
Mottled tablet appearance	Over-compaction of powder	Decrease compaction force
	Material sticking to rolls	See remedies for "sticking to compaction rolls"

Adapted from Smith T, et al. Development, scale-up, and optimization of process parameters: roller compaction. In: Developing solid oral dosage forms: pharmaceutical theory and practice. *Amsterdam: Elsevier; 2009. p. 715-24.*[123]

scale-up or scale-down a process. The data generated could be used to justify process parameters for an active formulation by taking into account how the active formulation's mechanical and physical properties differ from the placebo. In addition, this approach can be utilized to confer equivalence between equipment of the same scale. This type of assessment would be important when commercial production occurs on multiple roller compactors to ensure recipe set points are properly defined to ensure equivalent ribbon and granule properties.

32.8 TROUBLE-SHOOTING

A compilation of common roller compaction challenges and potential mitigation strategies are given in Table 32.15.

32.9 CONCLUSIONS

The roller compaction process is an effective and efficient means of enhancing powder behavior, in terms of flow and compactability, for downstream pharmaceutical processes. Over the past couple decades, roller compaction has become the predominant granulation method due to its many advantageous characteristics (eg, it does not require heat, solvents) and its popularity continues to increase, especially as the industry moves toward continuous production processes. This chapter has highlighted the benefits of roller compaction processes over alternate processing techniques, essential aspects of equipment, formulation, and process design, and critical considerations for scale-up, as well as reviewed both traditional and novel characterization tools, techniques, and models to facilitate efficient process development methods. Another important trend featured in this chapter was the growing use of small-scale tools and methods to simulate the roller compaction process, many of which can be conducted at bench-scale to enable efficient execution of roller compaction development processes, even at early stages of drug product development.

References

1. Cunningham JC. *Experimental studies and modelling of the roller compaction of pharmaceutical powders*. PhD Dissertation-Drexel University; 2005.
2. Guigon P, et al. *Roll pressing. Handbook of powder technology*. Oxford: Elsevier; 2007. p. 256–86..
3. Bindhumadhavan G. *Roller compaction of pharmaceutical powders*. PhD Dissertation- University of Birmingham; 2004.
4. Kuhn HA, Downey CL. Deformation characteristics and plasticity theory of sintered powder materials. *Int J Powder Met* 1971;**7**(1):15–25.
5. Dehont FR, et al. Briquetting and granulation by compaction new granulator-compactor for the pharmaceutical industry. *Drug Dev Ind Pharm* 1989;**15**(14):2245–63.
6. Adeyeye MC. Roller compaction and milling pharmaceutical unit processes: part I. *Am Pharm Rev* 2000;**3**:37–42.
7. Falzone AM, Peck GE, Mccabe GP. Effects of changes in roller compactor parameters on granulations produced by compaction. *Drug Dev Ind Pharm* 1992;**18**(4):469–89.
8. Yu L, et al. Understanding pharmaceutical quality by design. *AAPS J* 2014;**16**(4):771–83.
9. Leane M, Pitt K, Reynolds G. A proposal for a drug product Manufacturing Classification System (MCS) for oral solid dosage forms. *Pharm Dev Technol* 2015;**20**(1):12–21.
10. Rohrs BR, et al. Particle size limits to meet USP content uniformity criteria for tablets and capsules. *J Pharm Sci* 2006;**95**(5):1049–59.
11. Wells JI. Pharmaceutical preformulation: the physicochemical properties of drug substances. In: Rubinstein. MH, editor. *Pharmaceutical technology*. Chichester: Ellis Horwood Limited; 1988.
12. Sprockel OL, Stamato HJ. Design and scale-up of dry granulation processes. In: Ende DJA, editor. *Chemical engineering in the pharmaceutical industry: R&D to manufacturing*. Hoboken, NJ: John Wiley & Sons; 2011.
13. Gamble JF, et al. Roller compaction: application of an in-gap ribbon porosity calculation for the optimization of downstream granule flow and compactability characteristics. *Pharm Dev Technol* 2010;**15**(3):223–9.
14. Guigon P, Simon O. Roll press design-influence of force feed systems on compaction. *Powder Technol* 2003;**130**(1–3):41–8.
15. Lecompte T, et al. Dry granulation of organic powders—dependence of pressure 2D-distribution on different process parameters. *Chem Eng Sci* 2005;**60**(14):3933–40.

16. Miguelez-Moran AM, et al. Characterisation of density distributions in roller-compacted ribbons using micro-indentation and X-ray micro-computed tomography. *Eur J Pharm Biopharm* 2009;**72**(1):173–82.
17. Simon O, Guigon P. Correlation between powder-packing properties and roll press compact heterogeneity. *Powder Technol* 2003;**130**(1–3):257–64.
18. Han LH, et al. A modified Drucker–Prager Cap model for die compaction simulation of pharmaceutical powders. *Int J Solids Struct* 2008;**45**(10):3088–106.
19. Farber L, et al. Unified compaction curve model for tensile strength of tablets made by roller compaction and direct compression. *Int J Pharm* 2008;**346**(1–2):17–24.
20. He X, Secreast PJ, Amidon GE. Mechanistic study of the effect of roller compaction and lubricant on tablet mechanical strength. *J Pharm Sci* 2007;**96**(5):1342–55.
21. Hein S, Picker-Freyer KM, Langridge J. Simulation of roller compaction with subsequent tableting and characterization of lactose and microcrystalline cellulose. *Pharm Dev Technol* 2008;**13**(6):523–32.
22. Herting MG, Kleinebudde P. Roll compaction/dry granulation: effect of raw material particle size on granule and tablet properties. *Int J Pharm* 2007;**338**(1–2):110–18.
23. Malkowska S, et al. Effect of re-compression on the properties of tablets prepared by moist granulation. *Drug Dev Ind Pharm* 1983;**9**(3):349–61.
24. Sun C, Himmelspach MW. Reduced tabletability of roller compacted granules as a result of granule size enlargement. *J Pharm Sci* 2006;**95**(1):200–6.
25. Herting MG, Kleinebudde P. Studies on the reduction of tensile strength of tablets after roll compaction/dry granulation. *Eur J Pharm Biopharm* 2008;**70**(1):372–9.
26. Zinchuk AV, Mullarney MP, Hancock BC. Simulation of roller compaction using a laboratory scale compaction simulator. *Int J Pharm* 2004;**269**(2):403–15.
27. Kleinebudde P. Roll compaction/dry granulation: pharmaceutical applications. *Eur J Pharm Biopharm* 2004;**58**(2):317–26.
28. Miller RW. Roller compaction technology. In: Parikh D, editor. *Handbook of pharmaceutical granulation technology*. 2nd ed. New York, NY: Taylor & Francis; 2005. p. 159–72.
29. Lim H, et al. Assessment of the critical factors affecting the porosity of roller compacted ribbons and the feasibility of using NIR chemical imaging to evaluate the porosity distribution. *Int J Pharm* 2011;**410**(1–2):1–8.
30. Lubjuhn U, Sander U, Schonert K. Pressure profile in the compression zone of the high-pressure roller mill. *Zement Kalk Gips* 1994;**47**(4):192–9.
31. Wenerstrum S. Ten things to consider when choosing and installing a roller press system. *Powder Bulk Eng* 2000;**14**(2):37–47.
32. Dawes J, et al. Application of external lubrication during the roller compaction of adhesive pharmaceutical formulations. *Pharm Dev Technol* 2013;**18**(1):246–56.
33. Fanakoshi Y, Asogawa T, Satake E. Use of a novel roller compactor with concavo-convex roller pair to obtain uniform compacting pressure. *Drug Dev Ind Pharm* 1977;**6**(3):555–73.
34. Hlinak AJ, et al. Understanding critical material properties for solid dosage form design. *J Pharm Innov* 2006;**1**(1):12–17.
35. Cunninghham JC. *Experimental studies and modeling of the roller compaction of pharmaceutical powders. Materials engineering*. Philadelphia, PA: Drexel University; 2005. p. 267.
36. Katashinskii VP, Vinogradov GA, Kalutskii GY. Non-uniformity of strain and stress distribution over the width of rolled powder strip. *Poroshk Metall* 1975;**12**(156):28–32.
37. Guigon P, et al. Roll pressing. *Handb Powder Technol* 2007;**11**(Granulation):255–88.
38. McCann RJ. *Investigating the density distribution of roller compacted ribbons. Mechanical engineering*. Purdue University: UMI Dissertation Publishing; 2014.
39. Inghelbrecht S, et al. Instrumentation of a roll compactor and the evaluation of the parameter settings by neural networks. *Int J Pharm* 1997;**148**(1):103–15.
40. Bindhumadhavan G, et al. Roll compaction of a pharmaceutical excipient: experimental validation of rolling theory for granular solids. *Chem Eng Sci* 2005;**60**(14):3891–7.
41. Michrafy A, et al. Analysis of strain stress state in roller compaction process. *Powder Technol* 2011;**208**(2):417–22.
42. Nesarikar VV, et al. Instrumented roll technology for the design space development of roller compaction process. *Int J Pharm (Amsterdam, Neth.)* 2012;**426**(1–2):116–31.
43. Schonert K, Sander U. Shear stresses and material slip in high pressure roller mills. *Powder Technol* 2002;**122**:136–44.
44. Siebel E, Lueg W. Untersuchungen über die Spannungsverteilung im Walzenspalt. *Mitteilung der KWI* 1933;**15**:1–14.
45. Chekmarev AP, Klimenko PA, Vinogradov GA. Investigations of specific pressure, specific friction and the coefficient of friction during metal powder rolling. *Poroshk Metall* 1963;**2**(14):26–30. March-April, 1963.
46. Hatamura Y, Voneyama T. Measurements of actual stress and temperature on a roll surface during rolling. *JSME Int J* 1988;**31**:465–9.
47. Souihi N, et al. Roll compaction process modeling: transfer between equipment and impact of process parameters. *Int J Pharm* 2015;**484**:192–206.
48. Miguelez-Moran AM, Wu CY, Seville JPK. The effect of lubrication on density distributions of roller compacted ribbons. *Int J Pharm* 2008;**362**(1–2):52–9.
49. Patel BA, et al. Predicting the pressure distribution during roll compaction from uniaxial compaction measurements. *Chem Eng J (Amsterdam, Netherlands)* 2010;**164**(2–3):410–17.
50. Yu S, et al. A comparative study of roll compaction of free-flowing and cohesive pharmaceutical powders. *Int J Pharm (Amsterdam, Netherlands.)* 2012;**428**(1–2):39–47.
51. Johanson JR. A rolling theory for granular solids. *J Appl Mech* 1965;**32**(4):842–8.
52. Reynolds G, et al. Practical application of roller compaction process modeling. *Comput Chem Eng* 2010;**34**(7):1049–57.
53. Rowe JM, et al. Mechanistic insights into the scale-up of the roller compaction process: a practical and dimensionless approach. *J Pharm Sci* 2013;**102**(10):3586–95.
54. Shlieout G, et al. Dry granulation with a roller compactor part II: evaluating the operation modes. *Pharm Technol Eur* 2002;**14**(9):32–9.
55. Yusof YA, Smith AC, Briscoe BJ. Roll compaction of maize powder. *Chem Eng Sci* 2005;**60**(14):3919–31.
56. Miller RW. Roller compaction technology. In: Parikh DM, Parikh CK, editors. *Handbook on pharmaceutical granulation*. New York, NY: Marcel Dekker; 1997. p. 99–150.
57. Nesarikar VV, et al. Roller compaction process development and scale up using Johanson model calibrated with instrumented roll data. *Int J Pharm* 2012;**436**(1–2):486–507.
58. Inghelbrecht S, Paul Remon J. The roller compaction of different types of lactose. *Int J Pharm* 1998;**166**(2):135–44.
59. Inghelbrecht S, Remon JP. Roller compaction and tableting of microcrystalline cellulose/drug mixtures. *Int J Pharm* 1998;**161**(2):215–24.
60. Herting MG, Klose K, Kleinebudde P. Comparison of different dry binders for roll compaction/dry granulation. *Pharm Dev Technol* 2007;**12**(5):525–32.
61. Yu S, et al. The effects of lubrication on roll compaction, ribbon milling and tabletting. *Chem Eng Sci* 2013;**86**:9–18.

62. Bowden FP, Tabor D. *The friction and lubrication of solids*. Oxford, New York, NY: Oxford University Press; 1950.
63. Wu CY, et al. Roller compaction of moist pharmaceutical powders. *Int J Pharm* 2010;**391**(1–2):90–7.
64. Cunningham JC, Winstead D, Zavaliangos A. Understanding variation in roller compaction through finite element-based process modeling. *Comput Chem Eng* 2010;**34**(7):1058–71.
65. Dec RT, Zavaliangos A, Cunningham JC. Comparison of various modeling methods for analysis of powder compaction in roller press. *Powder Technol* 2003;**130**(1–3):265–71.
66. Gereg GW, Cappola ML. Roller compaction feasibility for new drug candidates. *Pharm Technol* 2002;**26**(Suppl):14–23.
67. Hilden J, Earle G, Lilly E. Prediction of roller compacted ribbon solid fraction for quality by design development. *Powder Technol* 2011;**213**(1–3):1–13.
68. Webb PA. *Volume and density definitions and determination methods*. Micromeritics Instrument Corporation; 2001. <http://www.micromeritics.com/Repository/Files/Volume_and_Density_determinations_for_Particle_Technologists_0.pdf>.
69. Iyer RM, et al. A novel approach to determine solid fraction using a laser-based direct volume measurement device. *Pharm Dev Technol* 2013;**19**(5):577–82.
70. Hancock BC, Clas S-D, Christensen K. Micro-scale measurement of the mechanical properties of compressed pharmaceutical powders. 1: the elasticity and fracture behavior of microcrystalline cellulose. *Int J Pharm* 2000;**209**(1–2):27–35.
71. Rowe RC, Roberts RJ. Mechanical properties. In: Alderborn G, Nystrom C, editors. *Pharmaceutical powder compaction technology*. New York, NY: Marcel Dekker; 1996. p. 283–322.
72. Stanley P. Mechanical strength testing of compacted powders. *Int J Pharm* 2001;**227**(1–2):27–38.
73. York P, et al. Fracture mechanics of microcrystalline cellulose powders. *Int J Pharm* 1990;**66**(1–3):143–8.
74. Fonteyne M, et al. Process analytical technology for continuous manufacturing of solid-dosage forms. *TrAC Trends Anal Chem* 2015;**67**(0):159–66.
75. Khorasani M, et al. Near-infrared chemical imaging (NIR-CI) as a process monitoring solution for a production line of roll compaction and tableting. *Eur J Pharm Biopharm* 2015;**93**(0):293–302.
76. Akseli I, et al. A quantitative correlation of the effect of density distributions in roller-compacted ribbons on the mechanical properties of tablets using ultrasonics and X-ray tomography. *AAPS PharmSciTech* 2011;**12**(3):834–53.
77. Gupta A, et al. Nondestructive measurements of the compact strength and the particle-size distribution after milling of roller compacted powders by near-infrared spectroscopy. *J Pharm Sci* 2004;**93**(4):1047–53.
78. Gupta A, et al. Influence of ambient moisture on the compaction behavior of microcrystalline cellulose powder undergoing uniaxial compression and roller-compaction: a comparative study using near-infrared spectroscopy. *J Pharm Sci* 2005;**94**(10):2301–13.
79. Gupta A, et al. Real-time near-infrared monitoring of content uniformity, moisture content, compact density, tensile strength, and young's modulus of roller compacted powder blends. *J Pharm Sci* 2005;**94**(7):1589–97.
80. Austin J, et al. The use of near-infrared and microwave resonance sensing to monitor a continuous roller compaction process. *J Pharm Sci* 2013;**102**(6):1895–904.
81. Acevedo D, et al. Evaluation of three approaches for real-time monitoring of roller compaction with near-infrared spectroscopy. *AAPS PharmSciTech* 2012;**13**(3):1005–12.
82. McAuliffe MAP, et al. The use of PAT and off-line methods for monitoring of roller compacted ribbon and granule properties with a view to continuous processing. *Org Process Res Dev* 2015;**19**(1):158–66.
83. Quyet PV, et al. A prediction model for monitoring ribbed roller compacted ribbons. *J Pharm Sci* 2013;**102**(8):2667–78.
84. Soh JP, et al. Importance of raw material attributes for modeling ribbon and granule properties in roller compaction: multivariate analysis on roll gap and NIR spectral slope as process critical control parameters. *J Pharm Innov* 2007;**2**(3–4):106–24.
85. Corredor CC, Bu D, Both D. Comparison of near infrared and microwave resonance sensors for at-line moisture determination in powders and tablets. *Anal Chim Acta* 2011;**696**(1–2):84–93.
86. Souihi N, et al. Near-infrared chemical imaging (NIR-CI) on roll compacted ribbons and tablets—multivariate mapping of physical and chemical properties. *Int J Pharm* 2015;**483**(1–2):200–11.
87. Shima S, Yamada M. Compaction of metal powder by rolling. *Powder Metall* 1984;**27**(1):39–44.
88. Slater RAC. *Engineering plasticity: theory and application to metal forming processes*. New York, NY: John Wiley & Sons, Inc; 1977.
89. Muliadi AR, Litster JD, Wassgren CR. Modeling the powder roll compaction process: comparison of 2-D finite element method and the rolling theory for granular solids (Johanson's model). *Powder Technol* 2012;**221**:90–100.
90. Martin CL, Bouvard D. Study of the cold compaction of composite powders by the discrete element method. *Acta Materialia* 2003;**51**:373–86.
91. Odagi K, Tanaka T, Tsuji Y. Compressive flow property of powder in roll-type presses-numerical simulation by discrete element method. *J Soc Powder Technol Jpn* 2001;**38**:150–9.
92. Esnault V, et al. Numerical simulation of roll compaction of aerated powders. *Chem Eng Sci* 2013;**104**(0):717–26.
93. Katashinskii VP. Analytical determination of specific pressure during the rolling of metal powders. *Poroshk Metall* 1966;**10**:1–10.
94. Katashinskii VP, Shtern MB. Stress-strained state of powder being rolled in the densification zone. I. Mathematical model of rolling in the densification zone. *Poroshk Metall* 1983;**11**(251):17–21.
95. Katashinskii VP, Shtern MB. Stress-strained state of powder being rolled in the densification zone. II distribution of density, longitudinal strain and contact stress in the densification zone. *Poroshk Metall* 1983;**12**(252):9–13.
96. Lee RS, Schwartz EG. An analysis of roll pressure distribution in powder rolling. *Int J Powder Met* 1967;**4S**:83–92.
97. Musikuin AM. Deformation region stresses in the cold rolling of metal powders. *Poroshk Metall* 1977;**12**(180):32–8.
98. Shima S, Oyane M. Plasticity theory for porous metals. *Int J Mech Sci* 1976;**18**(6):285–91.
99. Dec RT, Komarek RK. Computer aided design of roll type briquetters and compactors. *Proc Inst Briquet Agglom* 1991;**22**:1–14.
100. Dec RT, Komarek RK. Experimental investigations of roll press compaction. *Powder Handl Process* 1991;**4**(1):35–8.
101. Garner S, Strong J, Zavaliangos A. The extrapolation of the Drucker–Prager/Cap material parameters to low and high relative densities. *Powder Technol* 2015;**283**(0):210–26.
102. Michrafy A, et al. Experimental and numerical analyses of homogeneity over strip width in roll compaction. *Powder Technol* 2011;**206**(1–2):154–60.
103. Muliadi AR, Litster JD, Wassgren CR. Validation of 3-D finite element analysis for predicting the density distribution of roll compacted pharmaceutical powder. *Powder Technol* 2013;**237**(0):386–99.
104. Turkoglu M, et al. Modeling of a roller-compaction process using neural networks and genetic algorithms. *Eur J Pharm Biopharm* 1999;**48**(3):239–45.

105. Mansa RF, et al. Using intelligent software to predict the effects of formulation and processing parameters on roller compaction. *Powder Technol* 2008;**181**(2):217–25.
106. Jenike AW, Shield RT. On the plastic flow of Coulomb solids beyond original failure. *J Appl Mech* 1959;**81**:599–602.
107. Sommer K, Hauser G. Flow and compression properties of feed solids for roll-type presses and extrusion presses. *Powder Technol* 2003;**130**(1–3):272–6.
108. Dieter GE. *Mechanical metallurgy*. 3rd ed. New York, NY: McGraw-Hill Inc; 1986.
109. Balicki M. *Numerical methods for predicting roll press powder compaction parmaeters*. Albi: Ecole Des mines D'Albi-Carmaux; 2003.
110. Zienkiewicz OC, Taylor RL, Zhu JZ. *The finite element method: its basis and fundamentals*. 7th ed. Oxford: Elsevier; 2013.
111. Ketterhagen WR, Am Ende MT, Hancock BC. Process modeling in the pharmaceutical industry using the discrete element method. *J Pharm Sci* 2009;**98**(2):442–70.
112. LaMarche K, et al. Assessing materials' tablet compaction properties using the Drucker–Prager Cap model. *Powder Technol* 2014;**267**(0):208–20.
113. Cunningham JC, Sinka IC, Zavaliangos A. Analysis of tablet compaction. I. Characterization of mechanical behavior of powder and powder/tooling friction. *J Pharm Sci* 2004;**93**(8):2022–39.
114. Sinka IC, Cunningham JC, Zavaliangos A. Analysis of tablet compaction. II. Finite element analysis of density distributions in convex tablets. *J Pharm Sci* 2004;**93**(8):2040.
115. Pitt KG, Heasley MG. Determination of the tensile strength of elongated tablets. *Powder Technol* 2013;**238**:169–75.
116. Levin M. Pharmaceutical process scale-up. In: Levin. M, editor. *Drugs and the pharmaceutical sciences*, 3rd ed. vol. 118. New York, NY: Marcel Dekker, Inc; 2002.
117. Leuenberger H. Scale-up in the field of granulation and drying. In: Levin M, editor. *Pharmaceutical process scale-up*. New York, NY: Marcel Dekker, Inc; 2002. p. 151–70.
118. Hussain AS. A collaborative search for efficient methods of ensuring unchanged product quality and performance during scale-up of immediate-release solid oral dosage forms. In: Levin M, editor. *Pharmaceutical process scale-up*. New York, NY: Marcel Dekker, Inc; 2002. p. 325–52.
119. SUPAC-IR: immediate-release solid oral dosage forms: scale-up and post-approval changes: chemistry, manufacturing and controls, in vitro dissolution testing, and in vivo bioequivalence documentation, Food and Drug Administration Center for Drug Evaluation and Research, Editor. <http://www.fda.gov/cder/guidance/index.htm>; 1995.
120. SUPAC-IR/MR: immediate release and modified release solid oral dosage forms manufacturing equipment addendum, Food and Drug Administration Center for Drug Evaluation and Research, Editor. <http://www.fda.gov/cder/guidance/index.htm>; 1999.
121. Troup GM, Georgakis C. Process systems engineering tools in the pharmaceutical industry. *Comput Chem Eng* 2013;**51**:157–71.
122. Burke G. et al., Roller compaction process optimization using FBRM® at line particle characterization. Application notes. <http://us.mt.com/global/en/home/supportive_content/application_editorials/compaction.rxHgAwXLlLnPBMDSzq>; 2011.
123. Smith T, et al. *Development, scale-up, and optimization of process parameters: roller compaction. Developing solid oral dosage forms: pharmaceutical theory and practice*. Amsterdam: Elsevier; 2009. p. 715–24.

This page intentionally left blank

CHAPTER

33

Development, Optimization, and Scale-Up of Process Parameters: Tablet Compression

D. Natoli[1], *M. Levin*[2], *L. Tsygan*[2] *and L. Liu*[3]

[1]Natoli Engineering Company, St. Charles, MO, United States [2]Metropolitan Computing Corporation, East Hanover, NJ, United States [3]Pfizer Inc., Morristown, NJ, United States

33.1 INTRODUCTION

Compressing pharmaceutical tablets is one of the most efficient processes for producing a single dose of medication. Tablets are accepted and trusted by professionals and consumers alike; they are easily administered and simple to dose.

Compressing powders into a solid form dates back thousands of years. It wasn't until the early 1800s that tablet compression was automated in the sense that the hand crank was replaced by a leather belt and a steam-driven power bar. Early single-station tablet presses produced an average of 100 tablets per minute (TPM), while maintaining uniformity requirements for tablet hardness, thickness, and weight. Soon after, single-station presses were fading and making room for new technology: the rotary tablet press. Introduced in the mid-1800s, the rotary tablet press boasted speeds capable of compressing 640 TPM compared with high-speed presses today, capable of compressing up to 24,000 TPM. Although tablet presses have evolved through the years, the same principles apply (Fig. 33.1).

33.2 OPERATION PRINCIPLES OF COMPRESSION BY ROTARY PRESS

Modern rotary press usually include the following components:

Name of component	Function(s)
Hopper	Containing powder
Feeder	Distribute power into die cavity
Gravity, forced, centrifugal/die feed	Provide additional mixing
Turret	Position and progress the punch and die
Punch, lower and higher	Accept powder from feeder
	Contain powder until compression occurs
	Densify and compress the powder
	Eject the compressed tablet
Feed cam	Lower the lower punch to allow powder to come in
Compression cam, upper/lower	Force upper and lower punch to compress
Ejection cam	Lift the lower punch and push tablet above the turret
Ejection plate	Remove tablet off turret
Hydraulic	Provide hydraulic pressure
Instrumentation	Monitor force, speed, distance, etc.
Air handling	Provide exhaust

33.3 TOOL DESIGN

Good granulation is important for compressing quality tablets. If the granulation is poor, so will the long-term results. A proper tablet granulation will have good flow, compressibility, and release properties. Tablet compression tooling is equally responsible for the success of a tableting program. Tooling must

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00033-9

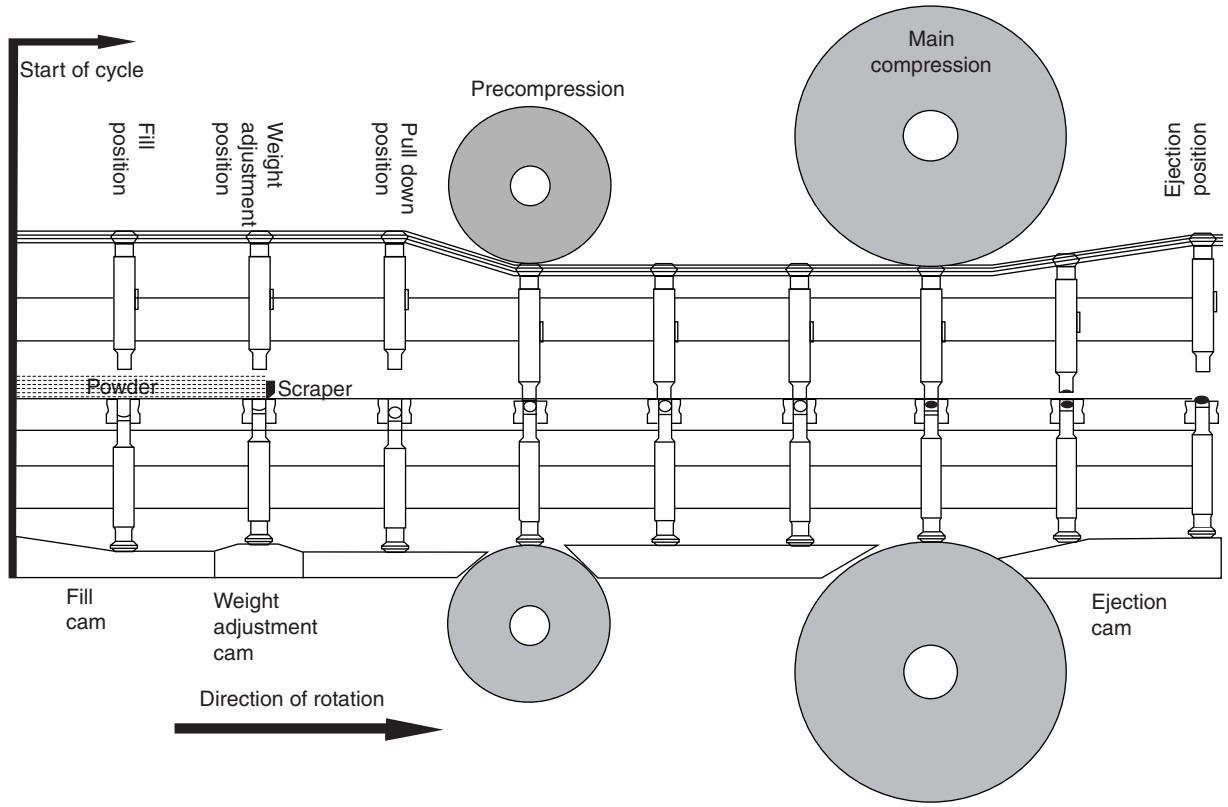


FIGURE 33.1 Rotary tablet press cycle.

be engineered to withstand the stresses associated with tablet compression, provide satisfactory service life, and maintain physical tablet uniformity. Proper tablet design is also critical. Pharmaceutical marketing departments feverishly attempt to design unique tablets, anticipating the design will quickly become branded and trusted in the eye of the consumer. A proper tool design is essential for placing that innovative design into the consumer's view.

The basic knowledge of tablet compression tooling and tablet design can save literally millions of dollars, prevent product loss, and reduce unnecessary equipment downtime. Understanding the basic physics of tablet compression will greatly enhance the ability to compress quality tablets more efficiently and provide better knowledge to troubleshoot and identify potential pitfalls before they happen (and they do!).

Communication is important with any tableting campaign. Marketing, research and design (R&D), engineering, production, and the tooling supplier must be in accordance and communicate new product-design and production requirements. The ideas and responsibilities of these departments may vary, but they share the common goal of manufacturing a quality tablet efficiently and productively.

33.3.1 Terminology

To communicate properly and understand the following material, it is important to have a basic understanding of the terminology used in this industry. Although these terms are most common and accepted, some terms may vary slightly between countries. This article deals with the terminology and general information related to the most common tablet shapes and rotary press tooling, the "TSM," "B," "D," and "Euronorm" 19- and 21-mm configurations (Tables 33.1 and 33.2 and Fig. 33.2).

33.3.2 Common tooling standards

There are two internationally recognized standards for tablet compression tooling, the "TSM" and the "EU" standards. Both "TSM" and "EU" standards identify the physical tool configuration for "B" and "D" type tablet compression tools, their critical dimensions, and associated tolerances assuring tablet quality and an efficient tablet press operation.

The "TSM" tooling standard is recognized in the Americas and is considered exclusive in the United States. "TSM" stands for the *Tablet Specification Manual*

TABLE 33.1 Common Tooling Terminology

Term	Definition
Tooling set	A complete set of punches and dies to accommodate all stations in a tablet press
Tooling station	The upper punch, lower punch, and die, which accommodate one station in a tablet press
Head	The largest diameter of a common punch that contacts the machines cams and accepts the pressure from the pressure rollers
Head flat	The flat portion of the head, which makes contact with the pressure rollers and determines the maximum dwell time for compression
Top head angle	Angle from the outside head diameter to the top head radius; it allows for sufficient head thickness and smoother camming
Top head radius	The radius on the top of the head, which blends the top head angle to the head flat. Some head configurations may consist of only the head radius without the head angle. This radius makes the initial contact with the pressure roll and allows a smoother transition into the compression cycle
Head back angle	Sometimes referred to as the inside head angle; located underneath the top head angle or the top head radius, which contacts the machine camming for vertical movement of the punch within the punch guides
Neck	Located below the head and provides clearance as the punch cycles through the machine cams
Barrel or shank	The vertical bearing surface of a punch, which makes contact with the punch guides in the machine turret for vertical guidance
Barrel chamfer	Chamfers at the ends of the punch barrel, eliminate outside corners
Barrel-to-stem radius	The radius that blends the punch barrel to the stem
Stem	The area from the barrel to the edge of the punch tip
Tip length	The straight portion of the punch stem
Tip straight	The section of the tip that extends from the tip relief to the end of the punch tip; it maintains the punch tip size tolerance
Land	The area between the edge of the punch cup and the outside diameter of the punch tip; this adds strength to the tip to reduce punch-tip fracturing
Tip face or cup	The portion of the punch tip that determines the contour of the tablet face; it includes the tablet embossing
Cup depth	The depth of the cup from the highest point of the tip edge to the lowest point of the cavity
Tip relief	The portion of the punch stem, which is undercut or made smaller than the punch tip straight; most common for lower punches to aid in reducing friction from the punch tip and die wall as the punch travels through the compression cycle; the area where the punch tip and relief meet must be sharp to scrape product from the die wall as the lower punch travels down for the fill cycle
Key	A projection normally made of mild steel, which protrudes above the surface of the punch barrel. It maintains alignment of the upper punch for reentry into the die; mandatory on upper punches with multiple tips and all tablet shapes other than round; commonly used with embossed round tablet shapes when rotation of the punch causes a condition known as double impression
Key position	The radial and height position of a key on the punch barrel; not found in all presses
Punch overall length	The total length of a punch, other than flat-face tablet configurations, that is normally a reference dimension, which consists of a combination of the working length and the cup depth dimensions
Working length	The dimension from the head flat to the lowest measurable point of the tip face, responsible for the consistency of the tablet's overall thickness
Anneal	A heat-treating process used on fragile punch tips to decrease the hardness of the punch cups reducing punch-tip fracturing
Bakelite tip relief	An undercut groove between the lower punch tip straight and the relief; it assures a sharp corner to assist in scraping product adhering to the die wall; normally a purchased option for lower punches
Major axis	The largest dimension of a shaped tablet
Minor axis	The smallest dimension of a shaped tablet
End radius	The radius on either end of a capsule or oval-shaped tablet
Side radius	The radius on either side of an oval- or modified-shaped tablet
Band	The center section of a tablet between the cup profiles: it is governed by a direct relationship of the die cavity profile
Compound cup	A cup profile which consist of two or more radii
Embossed	The raised identification on a tablet or a punch face; an embossed punch tip results in a debossed tablet
Debossed	The depressed identification on a tablet or a punch face; a debossed punch tip results in an embossed tablet

TABLE 33.2 Tablet Terminology

Term	Definition
Major axis	The largest dimension of a shaped tablet
Minor axis	The smallest dimension of a shaped tablet
End radius	The radius on either end of a capsule or oval-shaped tablet
Side radius	The radius on either side of an oval- or modified-shaped tablet
Band	The center section of tablet between the cup profiles: it is governed by a direct relationship of the die cavity profile
Compound cup	A cup profile which consist of two or more radii
Embossed	The raised identification on a tablet or a punch face: an embossed punch tip results in a debossed tablet
Debossed	The depressed identification on a tablet or a punch face: a debossed punch tip results in an embossed tablet

and is published, revised, and distributed by the American Pharmacist Association in Washington, DC. The "TSM" standards, once known as the "IPT" standards, were originally developed in 1968 by a committee consisting of major US pharmaceutical companies. Their motivation was an attempt to maintain standardization for "B" and "D" tablet compression tooling, which provides interchangeability between tablet presses. The "TSM" provides engineered drawings that are a valuable reference for troubleshooting and tool inspection. Today, the "TSM" committee consists of professionals from the tablet press, tooling, and tablet manufacturing industries. The "TSM" also includes useful information such as standard cup configurations for round tablets and a reference to common bisects for breaking tablets into multiple uniform dosages.

The "EU" tooling standard is internationally recognized and is more widely used than its counterpart, the "TSM" standard. "EU," which is the acronym for "Eurostandard" and "Euronorm," is considered the European standard for interchangeable "B" and "D" type compression tools. The "EU" standards were created by Trevor Higgins in an attempt to establish a tooling "norm" that provides tool interchangeability with the most common "B" and "D" type European tablet presses. The "EU" standard is printed and distributed by I Holland Ltd, Nottingham, UK.

33.3.3 EU, TSM, B, and D type punches

The "TSM" and "EU" standards manuals provide mechanical drawings and technical information for "B" and "D" type tools, which constitutes a majority of the tool configurations used today. The "B" type configuration has a nominal punch barrel diameter of 0.750″/19 mm. The "B" type has two different die sizes. The larger "B" dies have a diameter of 1.1875″ (30.16 mm) and the smaller "BB" dies have a 0.945″ (24 mm) diameter. The "D" type has a larger nominal barrel diameter of 1″ (25.4 mm) and a die diameter of 1.500″ (38.10 mm.). The "B" and "D" tool designation identifies the physical tooling size and was coined by engineer Frank J. Stokes in the late 1800s.

Stokes resided in Philadelphia, PA when he developed the first commercially available rotary tablet in the United States—the Stokes B1 Rotary. The B1 rotary press was extremely successful and most in demand by pharmaceutical companies nationwide. Mr Stokes, realizing the need for compressing larger and heavier tablets, developed the Stokes D3 rotary tablet press. The D3 tablet press uses slightly larger punches and dies, increasing the overall capacity to compress larger and heavier tablets.

During the second industrial revolution, Stokes expanded manufacturing capabilities and operated a facility in England for international distribution. Stokes soon became the world's leading tablet press manufacture and sold tablet presses and tooling in nearly every industrialized country. The designation "B" and "D" quickly became the international standard for identifying a tablet press capacity and a tool configuration, as it remains today.

At the brink of World War II, Stokes left England and focused all manufacturing activities in Pennsylvania. Stokes left behind trained engineers and qualified manufacturing personnel who soon realized the potential of the tablet press market and began manufacturing tablet presses and tooling under the name "Manesty." As a marketing strategy, Manesty reengineered the punches and tablet press cams to enhance tooling life and provide better performance. The Manesty punch is similar to the original Stokes design, but is exclusive to Manesty presses and not interchangeable with the more popular Stokes tablet presses. Manesty called their tablet presses the "Manesty B3B" and the larger "Manesty D3a"

Manesty soon became a major supplier in the compression equipment industry and successfully competed against Stokes in the global market. In the mid 1980s, the tablet press industry exploded and press manufactures were competing with tablet press output and innovation. Accommodating newer and high-speed tablet presses, the original Manesty tooling standard was refined to provide better interchangeability with the most common "B" and "D" tablet presses, identified by the "Eurostandard," often referred to as the "EU" standard and the "EU" norm.

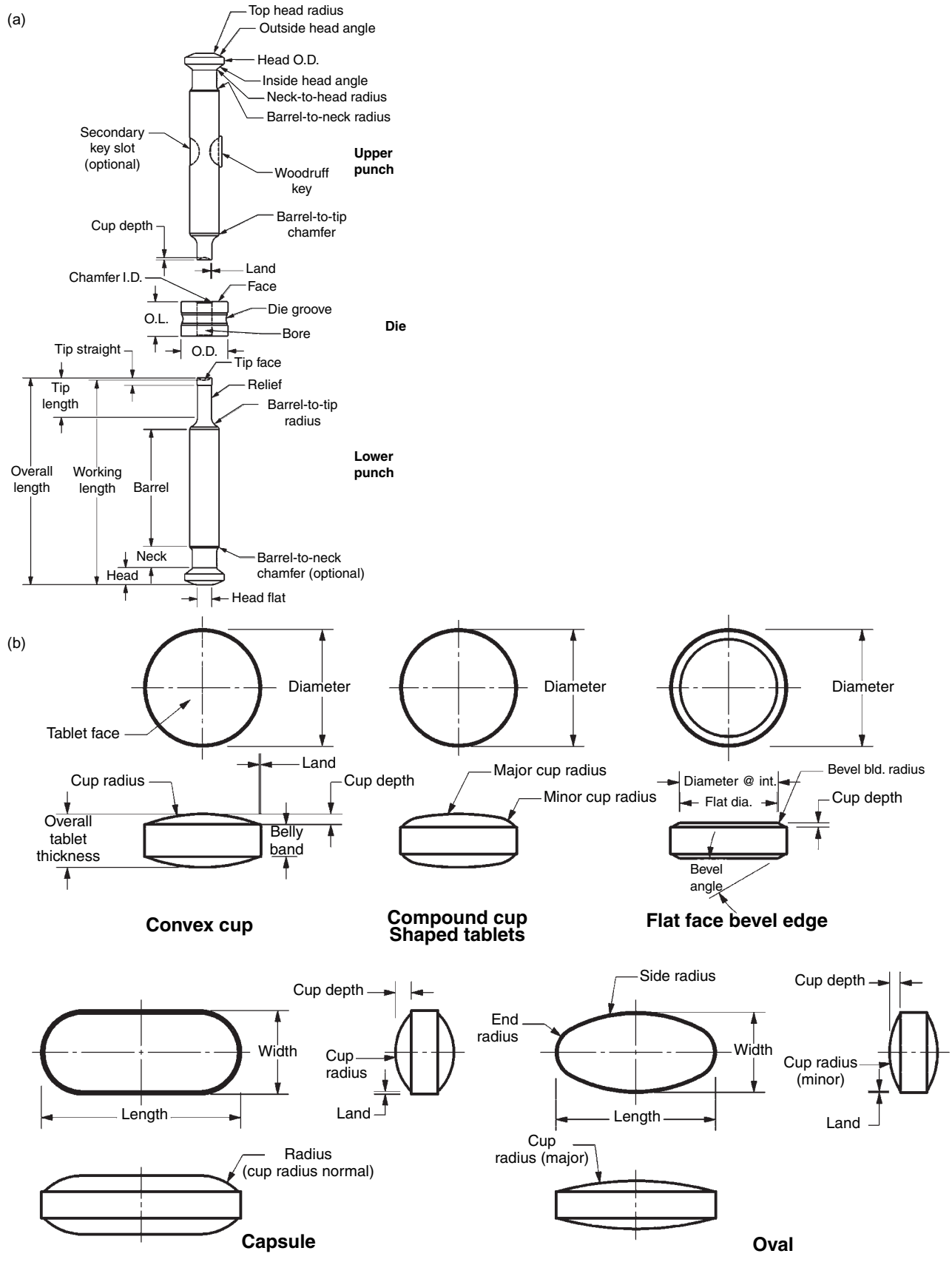


FIGURE 33.2 Punch and die terminology.

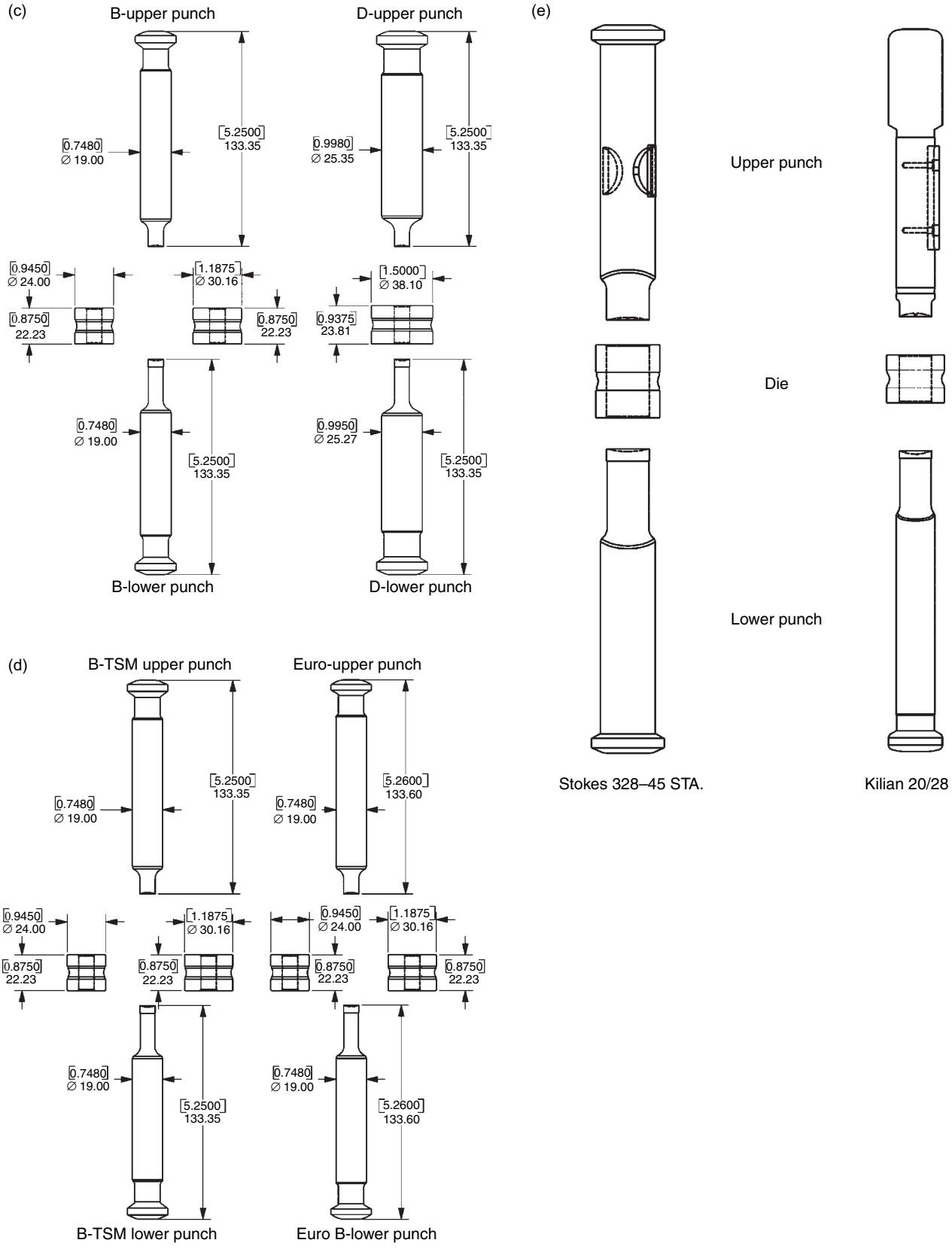


FIGURE 33.2 (Continued)

There are various models of tablet presses that do not conform to the standard "B" and "D" tool configurations and are engineered to be exclusive to a particular make and model of tablet press. Some of the more common configurations were designed in the early 1900s and are still used on tablet presses today. These unique tablet presses are generally larger and engineered to compress larger tablets more effectively. Kilian Gmbh, a division of IMA in Milan, Italy, is a major European manufacturer of tablet presses using the most common unique tool configuration. The Kilian-style upper punch does not use the common punch head configuration to guide the punches through the press cams; instead, the upper punch is guided by a machined cam angle located on the side of the upper barrel. The Kilian design provides a larger head flat, therefore, increasing the compression dwell time over the more popular "B" and "D" type tools.

33.3.4 Recent innovations

New technology continues to introduce innovative tool configurations in the effort to provide better efficiency of tablet press speed, product yield, cleaning, and safety.

In 1997, Ima introduced a line of unique tablet presses called the Ima Comprima. The Ima Comprima models use an innovative approach with tool design and granulation delivery. Unlike traditional tablet presses using a gravity-feed frame or force-feeding mechanism to fill the die with granulation, the Ima Comprima feeds the granulation through the die table, taking advantage of the centrifugal force created by the rotating turret for a rapid and uniform die fill. Unlike traditional presses, the Ima Comprima ejects the compressed tablet through the bottom of the die and uses gravity to eject the tablet from the press. Traditional tablet presses eject the tablet at the top of the die, requiring a mechanical stop or a take-off bar to physically contact and knock the tablet from the lower punch face. The Ima Comprima press is engineered to improve product yield, while providing a dust free environment for a more clean operation and a safer environment for the operator (Fig. 33.3).

The most recent innovation with tablet press and tooling technology is developed by Fette GmbH, located in Schwarzenbek, Germany. The new technology was introduced in 2005 and is favored by high-volume tablet manufactures. The technology does not use traditional compression dies, instead Fette developed die segments. Die segments provide an advantage over traditional dies by combining the tablet press turret die table and dies into three or five integral segments. Die segments are much easier and quicker to install than individual dies and die locks, reducing tablet press set-up time dramatically. More space is available around the turret circumference to increase the number of punches because the concept does not require the use of dies, which results in more tablets compressed per revolution than traditional presses of the same size.

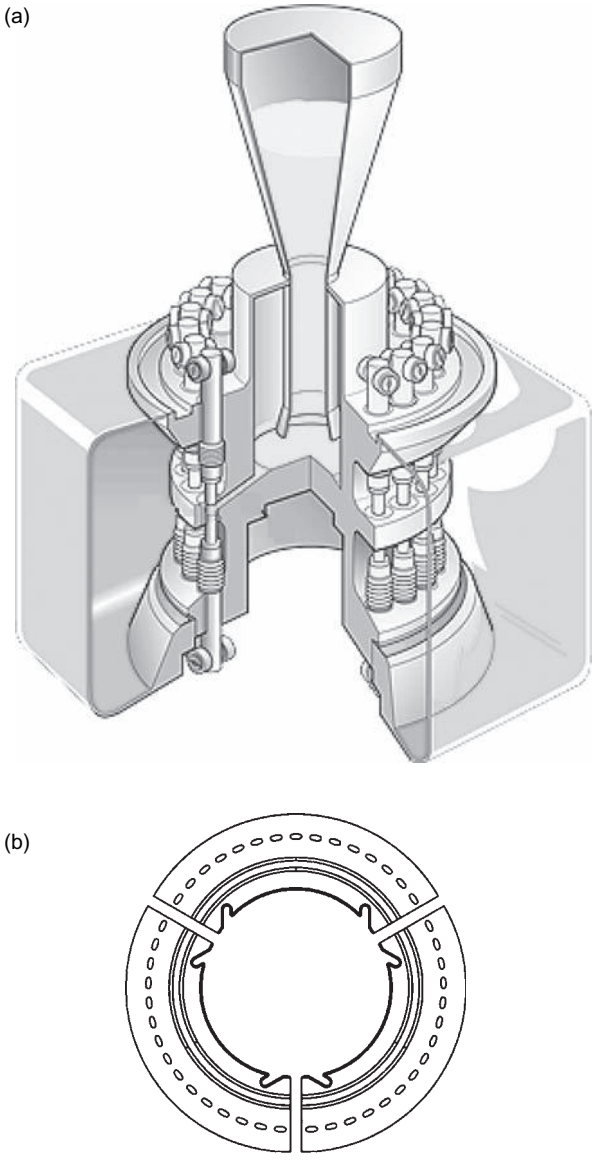


FIGURE 33.3 (a) IMA press and tools. (b) Fette 2090 die segments.

Tablet press technology has recently brought attention to the steel used for punches and dies with "wash in place" tablet presses. "Wash in place" tablet presses are becoming more common and are available

from most major tablet press suppliers. To reduce the possibilities of tool discoloration and corrosion, it is important that the tools are immediately removed and dried if the tools cannot be confirmed dry in the tablet press turret.

33.3.5 Cup depth, overall length, and working length

These are the most critical dimensions in any tooling program that relate directly to final tablet thickness, weight, and hardness. The overall length (O.L.), is a reference dimension and, therefore, does not have a specified tolerance. A reference dimension is defined by the Machinery Handbook (2) as

> A dimension, usually without a tolerance and used for information purpose only. It is considered auxiliary information and does not govern production or inspection operations. A reference dimension is the repeat of a dimension or is derived from other values shown on the drawing or on related drawings.

The two dimensions comprising the punch overall lengths are the working length and the cup depth, with the exception of flat-face tip configuration, which does not have a cup and is used to compress a wafer-type tablet. The two dimensions are the working length dimension, with a tolerance of plus or minus 0.001″, and the cup depth, with a tolerance plus or minus 0.003″. Combining the two tolerances that affect the O.L. of a punch, the calculated tolerance would be plus or minus 0.004″. The major concern with these dimensions is to maintain consistency within a set of punches in order to maintain tablet weight, hardness, and thickness. The more critical of the two dimensions is the working length (W.L.). The W.L. needs to be inspected as a single dimension and preferably for consistency within the given working-length tolerance, and not for a number formulated from the cup depth subtracted from the overall length. A set of punches should be separated into uppers and lowers and inspected for variances as such. For example, all of the upper punches are checked for length consistency, and then all of the lower punches are checked as a separate unit. As long as both upper and lower punches fall within the desired tolerance range, tablet thickness, hardness, and weight will be consistent.

Although the cup depth is not responsible for tablet thickness, it should be confirmed within the given tolerance to maintain tablet overall consistency, and it too should be inspected as single dimension (Fig. 33.4).

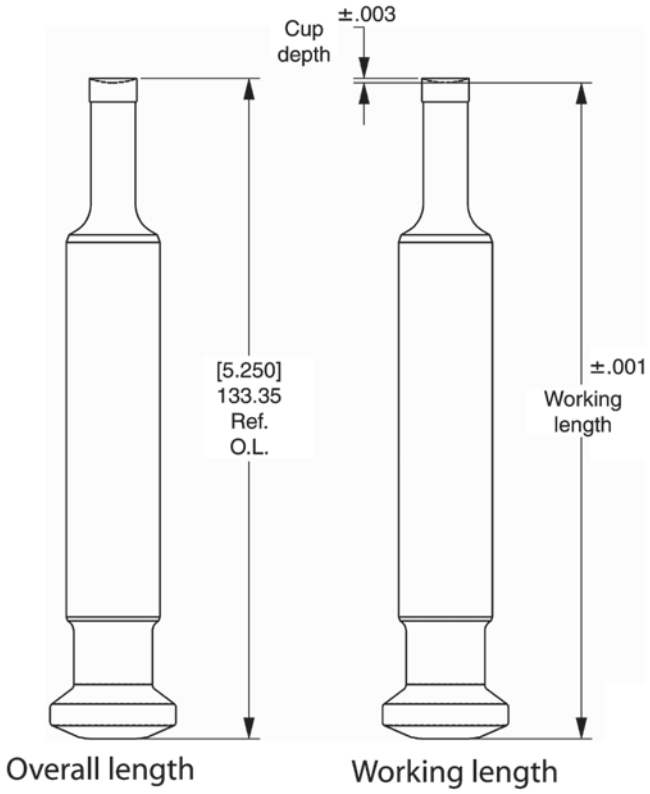


FIGURE 33.4 Overall and working lengths.

33.3.6 Tooling options

During the 1980s, the tablet compression industry was introduced to higher speed and more automated tablet presses that assured interchangeability with the "TSM" standard tool configurations. Although the standard tool configuration may be compatible, in some cases, it was not optimal and required minor modification to achieve expected performance. In addition, the standard tool configuration may not be desirable for compressing certain products. All products are different and have unique characteristics, and therefore, may require slight tooling modifications. Tablet manufactures need to be informed of available options to achieve the best possible performance from the tablet press and tooling. The next section provides a description of tooling options that can be a benefit on both high-speed and standard presses.

33.3.6.1 Common tooling options

33.3.6.1.1 Domed heads

The domed head configuration is adaptable to both the upper and lower punch, and maintains the identical top head radius and head flat as the "Eurostandard." It is an option only for the "TSM" head configuration and is compatible with the American "TSM" cams, and

should be considered for all high-speed tablet presses. As the speed of the tablet press continues to increase, tablet manufactures are coming to realize the advantage of the domed-style head with the larger top radius. The domed head style has several advantages over the standard "TSM" head profile. The larger 5/8″ radius on the domed head reduces the enormous stress that is more common with the smaller 5/16″ radius on a standard head when the punch makes initial contact with the pressure roller. This stress can cause a condition called head pitting, which is identified by voids on the head flat. The impact of the pressure roller and head radius at high speeds and heavy forces can cause a work-hardening effect, contributing to the pitting of the head flat. This form of pitting is detrimental to the life of the punches and pressure rollers. The domed head configuration provides a smoother transition into the compression cycle of the tablet press, reducing stress and premature wear of the pressure rollers (Fig. 33.5).

33.3.6.1.2 Punch head flats

The punch head flat is located at the opposite end of the punch tip. It is usually the flat on top of the largest diameter of the punch. The head flat provides dwell in the cycle of compression. Dwell as related to tablet compression can be defined as "time spent in the same position." "Time spent in the same position" refers to the time the punch is held vertically in-between the pressure rollers; this time is generally measured in milliseconds.

Dwell time is dictated by the rotational speed and pitch diameter of the turret and the size or diameter of the punch head flat. The punch head flat is an "important" dimension, but is not considered a critical dimension of the punch. As with all important dimensions of the punch, it is essential to maintain a dimensional consistency within the entire set. In general, the majority of powders or granulate formulated for compression will produce quality tablets that meet all requirements, even if the punch head flat diameter is spread and using the entire tolerance range. A discrepancy related to physical tablet specifications, which may be correlated with the inconsistency of the head flat, is typically uncommon. If maintaining tablet consistency becomes challenging, knowing and understanding the importance of the head flat becomes essential.

For products requiring longer dwell time, the punch option commonly referred to as the "extended head flat" is an attractive alternative to the standard head flat (Fig. 33.6a). The extended head flat option maximizes the flat area by reducing punch head top radii, thus resulting in a larger flat diameter. The extended head flat will increase the dwell time, or the time under full compression between the pressure rollers. The extended dwell time can aid in reducing tablet defects, such as low hardness, laminating, and even capping. The extended head flat option can also help to reach the maximum press speed, yielding optimal production for powders difficult to compress.

However, the extended head flat option does increase the physical head flat diameter to a dimension critically close to the punch neck diameter. If the maximum compression force or dwell initiates at a point on the top of the punch head, which is outside the neck diameter below, it can create dangerous stresses in in

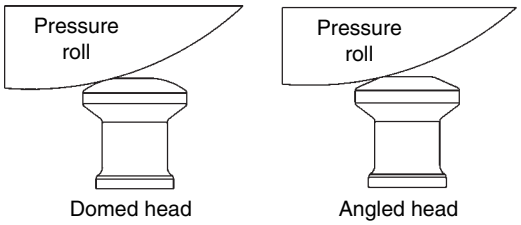


FIGURE 33.5 Note: Differences between TSM and TSM domed.

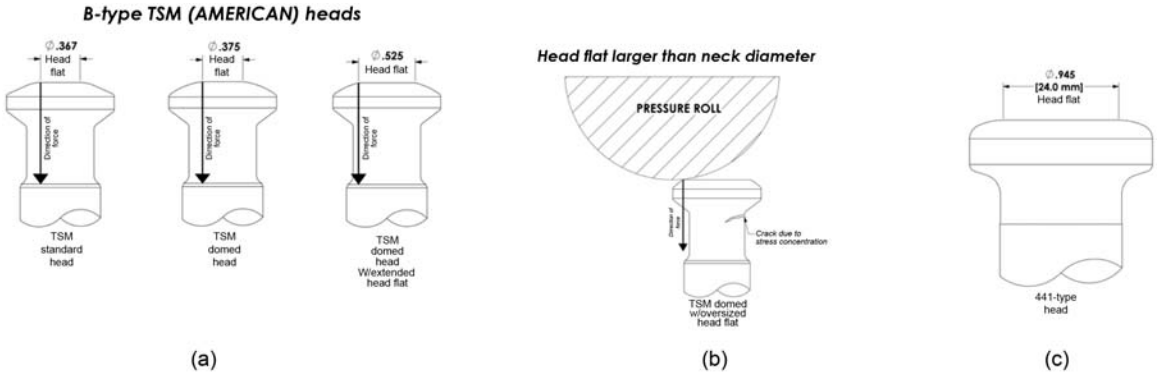


FIGURE 33.6 Dwell time.

the steel that are prone to fracture the punch head or even the punch neck, which may result in extensive machine damage (Fig. 33.6b).

Some tablet press manufactures provide an option to substantially increase the punch head flat dimension by using special cams and neckless style punches. This configuration is sometimes referred as the Neckless 441 design and is an option only for "D" type punches (Fig. 33.6c).

The procedure of reworking the punch head is common and is performed to reduce the rolling fatigue or pitting generally caused from excessive pressure, improper lubricants, and insufficient lubrication, contamination from airborne powders and improperly wearing and/or aligned machine parts. The procedure of reworking the head should only be performed by a trained technician in order to maintain the proper dimensions.

Introducing a new and unused spare punch into a well-worn set of punches should be done with caution because the new punch head flat may be inconsistent with the worn punch. It is advisable to check the critical working length and compare it with the worn punch to assure compatibility before adding it to the set.

33.3.6.1.3 Rotating heads

The rotating punch head is a two-part punch configuration, the head is separate from the punch barrel and tip, allowing the head to be removed and replaced as the head wears. When compressing round tablets, the punches will rotate as they are pulled around the cam track through the various stages of the tablet compression. As the punches rotate, the wear and stresses on the back angle of the head are distributed around the entire back angle-bearing surface. When compressing tablet shapes other than round, the punches do not rotate, causing the wear to be concentrated at a single point, which results in premature head wear. The wear is distributed along the entire surface of the back angle because the rotating head configuration allows the head to rotate when compressing nonround tablet shapes. This helps to decrease head wear and prolong the life of the punches (Fig. 33.7a).

33.3.6.1.4 Mirror-finished heads

Some high-speed tablet presses use heavy metal cams, such as bronze and bronze alloys. This material is good for eliminating premature head wear and prolonging tool life, but it has a negative effect because it contaminates the lubrication and turns it a black, dark-green color. The typical finish of a punch head is done with fine emery or fine abrasive pads. This finish leaves fine radial lines on the contact surfaces of the heads and has a filing effect on the softer cams,

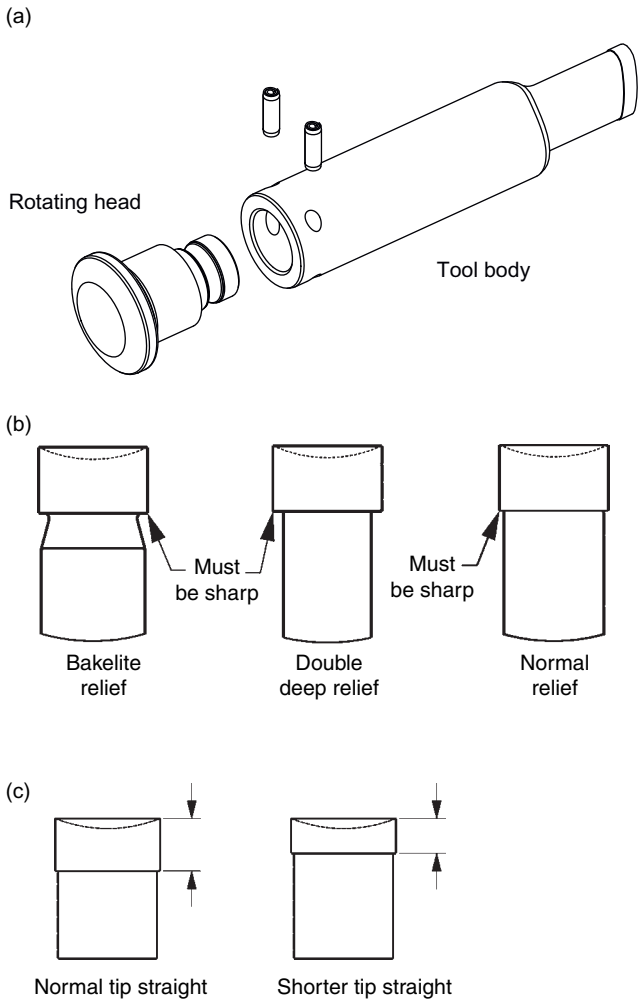


FIGURE 33.7 (a) Rotating head. (b) Punch tip reliefs. (c) Punch tip straights.

causing discoloration of the lubrication and premature cam wear. Polishing the punch heads with a soft cotton wheel and fine polishing compound to a mirror finish helps keep the lubrication cleaner and prolongs cam life.

33.3.6.1.5 Bakelite relief and double-deep relief

It is important to maintain a sharp edge around the lower punch tip relief. A sharp edge assists with the pull-down cycle of the lower punch after tablet ejection. If residual product is adhered to the die wall, the sharp lower punch tip relief will help scrape the die clean as well as cutting through the product to reduce the possibility of product wedged and re-compressed between the punch tip and die wall. Product wedged between the punch tip and die wall may cause excessive heat and thermal expansion of the punch tip. This could result in punch binding and/or seizure, premature head wear, tablet

discoloration, or burning and dark specs contaminating the tablet. A bakelite relief assures a sharp edge to assist with removing product adhered to the die wall, allowing the punch tip to move freely in the die. A "double deep relief" increases the depth of the lower punch relief and provides the same results as the bakelite relief; both designs assure a sharp edge at the punch tip. The bakelite relief is an added-cost option for punches, whereas the double-deep relief is generally a no charge option (Fig. 33.7b).

33.3.6.1.6 Short lower punch tip straight

The lower punch tip creates a tremendous amount of friction because it travels the full length of the die through the various stages of tablet compression. When compressing sticky products or products with a low melting point, the friction created by the lower punch tip can cause lower punch binding. Reducing the bearing surface of the lower punch tip will reduce friction, allowing the punch to travel easier in the die and reduce operating temperatures.

33.3.6.1.7 Punch-barrel chamfers

Punch-barrel chamfers are required on punches used with presses fitted with rubber or plastic guide seals. The barrel chamfer has an advantage over the common break edge for these press models. The absence of a chamfer on the tip end of the punch can create difficulties while installing punches. Forcing the punch past the seal can damage to the seals, resulting in seepage of lubrication from the upper-punch guides, inherently causing product contamination. Damaged lower guide seals can allow product seepage into the lower-punch guides as well as mixing with the lubrication, causing tight punches and possibly press seizure. A barrel chamfer on the head end of the punch can reduce wear of the punch guides caused by the punches being cocked from the torque of rotation as the punch travels vertically in the guides (Fig. 33.8).

33.3.6.1.8 Key types and positions

Punch barrel keys are mandatory for upper punches when compressing nonround tablets. The upper punch key maintains alignment of the tip for reentry into the die for compression. Keys are not generally required for lower punches because the lowers do not leave the die during the compression cycle, so maintaining alignment is not required. Keys may also be required when compressing round tablets with embossing to eliminate the punch from spinning after compression, causing damage to the embossed tablet and reducing the likelihood of a "double impression" on the tablet face. The punches may also require keys when the orientation of embossing for the top and bottom of the tablet is required to be constant.

Keys fitted to the upper punches are available in two configurations: (1) the standard Woodruff key, sometimes referred to as the pressed-in key and (2) the feather or flat key, often referred as the European key.

The Woodruff key, often referred to as the half-moon key because of its shape, is available in two styles, standard and the Hi-Pro. The Hi-Pro key has a tab on each side of the exposed top section and rests on the barrel. The taps keep the key secure by eliminating the rocking action common to the standard Woodruff. To obtain maximum security for high-speed presses, the Woodruff key is fastened into the barrel using screws. The Woodruff key can swell the barrel at the position of the key slot as is pressed into position, causing excessive drag and sometimes galling of the upper punch and punch guide.

The feather key is a longer, flat key, and comes in a variety of lengths, depending on the tablet press. Unlike the pressed in Woodruff key, the feather key fits into a milled slot and is secured into position using machine screws.

The height and radial position of a key is critical to obtain maximum press performance. Unfortunately, no standard has been established due to the particular requirements of the many styles of tablet presses. If the key is placed too low or is too long, it can interfere with the upper punch guide seal and cause damage and/or seepage of lubrication, resulting in product contamination. If the key is too high, it can travel out of the key slot at the top of the punch guide, resulting in severe damage to the punches and press (Fig. 33.9).

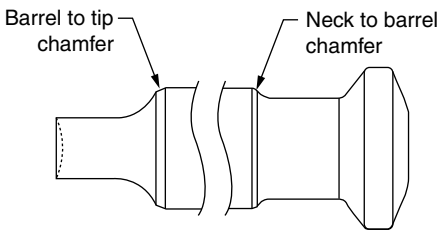


FIGURE 33.8 Punch and barrel chamfers.

33.3.7 Tool configuration for small and micro tablets

It is common to experience difficulties maintaining tablet hardness, thickness, and weight while compressing small and micro tablets. Compression force is sensitive and will generally require minimum forces. In some cases, the tablet is compressed by the weight of

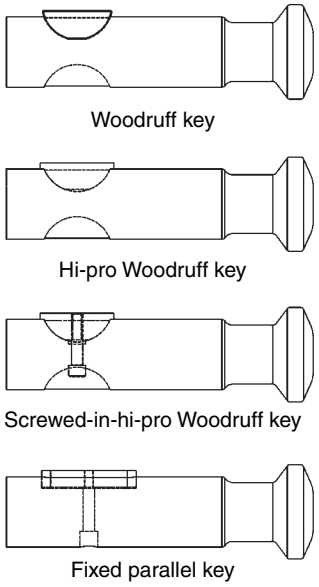


FIGURE 33.9 Key types.

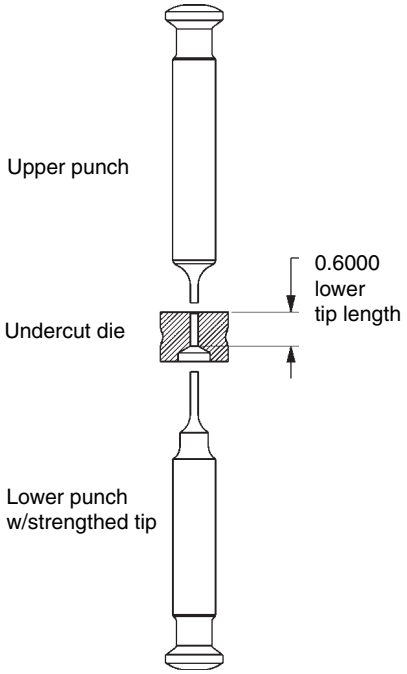


FIGURE 33.10 Tool configuration and micro tablets.

the punch. Excessive tonnage can distort the punch tip and alter the critical working length, making tablet consistency virtually impossible. Tip breakage is also frequent and can damage additional punches and the tablet press, most commonly the feed frame. A special tool configuration is recommended for compressing tablets smaller than 0.125″ (3 mm). This configuration modifies the punches and dies and is used in conjunction with a shallow fill cam that is fitted on the press to minimize lower punch travel in the die. The punch modification involves shortening the punch tips and eliminating the lower punch relief. Shortening the tip straights to their minimum length will strengthen the tip, considerably increasing the maximum compression force. The lower punch tip relief is removed to reduce the clearance between the tip stem and the die bore, providing additional support to the tip stem, thus decreasing distortion. Reducing the tip length increases the barrel length; therefore the bottom of the die is undercut to accept the longer barrel for tablet ejection (Fig. 33.10).

33.3.8 Tapered dies

A tapered die has numerous advantages. A die can be tapered on one side or on both sides, with the advantage of turning the die over and doubling its life. The biggest advantage of a tapered die is to exhaust trapped air in the product as the upper punch enters the die at the beginning of the compression cycle. This is especially helpful for deep-cup punches, fluffy granulation, and high-speed presses. A tapered die provides the ability to compress a harder tablet with the same amount of pressure required with a straight die. It is helpful in reducing capping and laminating. Taper will allow the tablet to expand at a slower rate because it is being ejected from the die, reducing stress that can cause lamination and capping. Taper decreases the ejection force, prolonging the life of the lower punch heads and ejection cam, thus reducing friction and allowing the press to operate at a lower temperature. Tapered dies help align the upper-punch tip upon entering the die, eliminating premature tip wear—this is especially helpful for presses with worn upper-punch guides. A standard taper on a “BB” or “D” die is 0.003″ by 3/16″ deep. Die taper can be tailored to meet special requirements. Although there are numerous advantages with using taper, there are also disadvantages. The upper punch can wedge in-between the punch tip and die wall as it’s pressed into the die because the taper is conical with the largest area at the top. Excess product can migrate between the punch tip and die bore due to the additional punch tip to die bore clearance as a result of the taper. If the upper punch is wedged and sticks in the die, it will be evident by spotty tablets and/or premature wear at the back angle of the upper punch (Fig. 33.11).

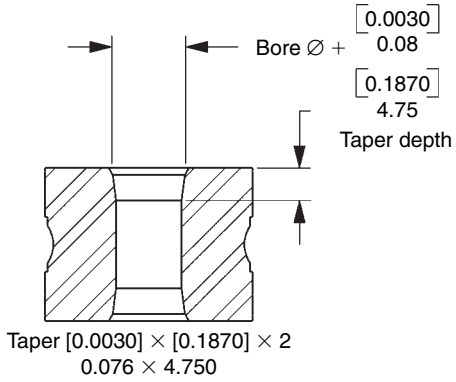


FIGURE 33.11 Tapered die.

33.4 TABLET DESIGNS

Proper punch face contour is essential for tooling life and tablet quality. The compression force should be determined during the R&D phase of a new product. If heavy compaction forces are required, a shallow- or standard-cup configuration should be considered to assure satisfactory tooling life and tablet quality. If the compaction force is to remain light to standard, a variety of configurations may be considered. Compression force has a lateral force that can expand the sides of the punch cup outward toward the die wall.

Excessive pressure can permanently distort and cause premature failure of the punch tip. For a high compaction force the cup may be strengthened by:

1. Increasing the land area on the punch tip to provide additional strength
2. Reducing the hardness of the punch tip, allowing the tip to flex without breaking
3. Increasing the cup radius or decreasing the cup depth to eliminate the damaging effect of flexing and abrasion to the inside of the cup

The flat-face bevel edge (FFBE) tablet configuration is subjected to the same lateral force. These edges can be strengthened by steps 1 and 2 and by increasing the radius between the flat and the bevel, which is normally 0.010–0.015″. The flat-face radius-edge (FFRE) configuration provides a stronger punch tip than the FFBE and can eliminate edge chipping by reducing sharp corners on the tablet face. Another common cup configuration is the compound cup. The compound cup has two radii, which makes the tablet roll better during the coating process, eliminating tablet-edge erosion. The compound cup design generally has more cup volume and is the optimum tablet design for heavy tablets because it generally reduces the tablet band, giving the tablet a thinner appearance. However, the compound cup is one of the weakest tablet designs

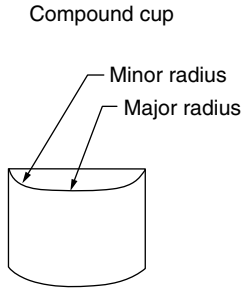


FIGURE 33.12 Compound cup.

due to the stresses created at the intersection of the two cup radii and the steep cup, which causes excessive abrasion during compression, shortening the tool life (Fig. 33.12).

Elaborate three-dimensional cup configurations are becoming more common in the candy and vitamin industry. It is critical that compaction forces are determined during the R&D phase and results provided to the tooling manufacture because of the high- and low-cup designs.

The concavity standards for round punch tips are published in the TSM. These standards (Table 33.3) include cup depths for shallow, standard, deep, extra deep, modified ball, FFBE, and FFRE. For radius cup designs, the "TSM" identifies the cup by the cup depth, whereas the European tableting industry identifies the cup by the cup radius (Fig. 33.13).

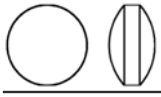
33.4.1 Tablet shapes

There are as many tablet shapes as there are applications, which are endless. Tablets are used in automobile airbags, batteries, soaps, fertilizers, desiccants, and buttons, just to name a few. Historically, round tablets were the most common, uncomplicated, and easy to set-up and maintain. Special-shape tablets are tablet shapes other than round, and include shapes such as capsule, oval, square, triangle, etc. Exotic shape tablets are more unique than round or special shapes. Exotic-shaped tablets include animal and heart-shaped tablets and other unique tablet shapes that require an internal radii or angle (Fig. 33.14).

A unique tablet shape will provide better tablet identification, which helps maintain consumer interest and loyalty.

The most common special shapes in the pharmaceutical industry are the capsule, modified capsule, and oval shapes. These shapes typically accommodate more volume and are more unique than standard rounds. A film-coated tablet is better to use with a modified capsule, rather than a capsule shape, to eliminate twinning during the coating process. A modified

TABLE 33.3 Examples of Tooling Standards

Tablet diameter						
Inches [mm]	**Shallow cup depth**	**Standard cup depth**	**Deep cup depth**	**Extra-deep cup death**	**Modified ball cup depth**	**FFBE/FFRE cup depth**
1/8 [3.175]	0.005 [0.127]	0.017 [0.432]	0.024 [0.610]	0.030 [0.762]	0.040 [1.016]	0.007 [0.178]
5/32 [3.970]	0.007 [0.178]	0.021 [0.533]	0.030 [0.762]	0.036 [0.914]	0.049 [1.245]	0.008 [0.203]
3/16 [4.763]	0.008 [0.203]	0.026 [0.635]	0.036 [0.914]	0.042 [1.067]	0.059 [1.499]	0.009 [0.229]
7/32 [5.555]	0.009 [0.229]	0.029 [0.737]	0.042 [1.067]	0.048 [1.219]	0.069 [1.753]	0.010 [0.254]
1/4 [6.350]	0.010 [0.254]	0.031 [0.787]	0.045 [1.143]	0.050 [1.270]	0.079 [2.007]	0.011 [0.279]
9/32 [7.142]	0.012 [0.305]	0.033 [0.838]	0.046 [1.161]	0.054 [1.372]	0.089 [2.261]	0.012 [0.305]
5/16 [7.938]	0.013 [0.330]	0.034 [0.584]	0.047 [1.194]	0.060 [1.524]	0.099 [2.515]	0.013 [0.330]
11/32 [8.730]	0.014 [0.356]	0.035 [0.889]	0.049 [1.245]	0.066 [1.676]	0.109 [2.769]	0.014 [0.356]
3/8 [9.525]	0.016 [0.406]	0.036 [0.914]	0.050 [1.270]	0.072 [1.829]	0.119 [3.023]	0.015 [0.381]
13/32 [10.318]	0.017 [0.432]	0.038 [0.965]	0.052 [1.321]	0.078 [1.981]	0.128 [3.251]	0.016 [0.406]
7/16 [11.113]	0.018 [0.457]	0.040 [1.016]	0.054 [1.372]	0.084 [2.134]	0.133 [3.378]	0.016 [0.406]
15/32 [11.905]	0.020 [0.508]	0.041 [1.041]	0.056 [1.422]	0.090 [2.286]	0.148 [3.759]	0.016 [0.406]
1/2 [12.700]	0.021 [0.533]	0.043 [1.092]	0.056 [1.499]	0.095 [2.413]	0.158 [4.013]	0.016 [0.406]
17/32 [13.493]	0.022 [0.559]	0.045 [1.143]	0.061 [1.549]	0.101 [2.565]	0.168 [4.267]	0.016 [0.406]
9/16 [14.288]	0.024 [0.610]	0.046 [1.168]	0.063 [1.600]	0.107 [2.718]	0.178 [4.521]	0.016 [0.406]
19/32 [15.080]	0.025 [0.625]	0.048 [1.219]	0.066 [1.676]	0.113 [2.870]	0.188 [4.775]	0.016 [0.406]
5/8 [15.875]	0.026 [0.660]	0.050 [1.270]	0.065 [1.727]	0.119 [3.023]	0.198 [5.029]	0.016 [0.406]
11/16 [17.483]	0.029 [0.737]	0.054 [1.372]	0.073 [1.854]	0.131 [3.327]	0.217 [5.512]	0.020 [0.508]
3/4 [19.050]	0.031 [0.787]	0.058 [1.473]	0.078 [1.981]	0.143 [3.632]	0.237 [6.020]	0.020 [0.508]
13/16 [20.638]	0.034 [0.864]	0.061 [1.549]	0.083 [2.108]	0.155 [3.937]	0.257 [6.528]	0.020 [0.508]
7/8 [22.225]	0.037 [0.940]	0.065 [1.651]	0.089 [2.260]	0.167 [4.242]	0.277 [7.036]	0.020 [0.508]
15/16 [23.813]	0.039 [0.991]	0.069 [1.753]	0.094 [2.388]	0.179 [4.547]	0.296 [7.518]	0.020 [0.508]
1 [25.400]	0.042 [1.067]	0.073 [1.854]	0.099 [2.515]	0.191 [4.851]	0.316 [8.026]	0.025 [0.635]

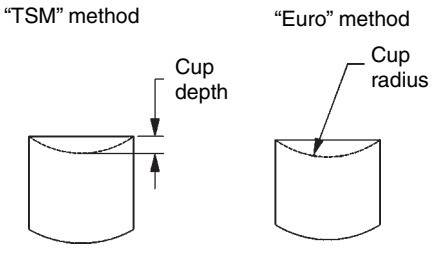


FIGURE 33.13 For radius cup designs, the TSM identifies the cup by the cup depth, whereas the European tableting industry identifies the cup by the cup radius.

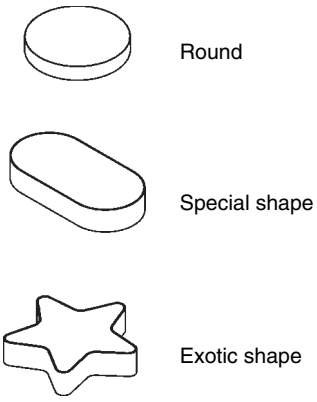


FIGURE 33.14 Basic tablet shapes.

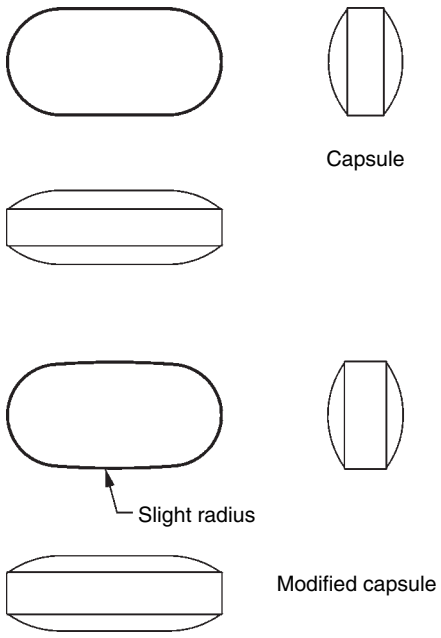


FIGURE 33.15 Capsule and modified capsule shapes.

capsule shape can be designed to have the appearance of a capsule shape with the advantage of a radius on the major axis, reducing the contact surface area during the coating process (Fig. 33.15).

33.4.2 Tablet face configurations

Tablet shapes are virtually infinite, as are tablet face configurations. The tablet face configuration is commonly referred to as the "cup" of the punch. The cup is the area at the tip end of the punch that is responsible for the configuration of the top and bottom of a tablet. The "TSM" provides cup depth standards for the six most common cup configurations for round tablets.

The "TSM" defines the cup depth of single radius tablet configurations by the depth of the concavity and differs from the "EU" configurations, which use the cup radius value. The cup radius is more difficult to check and to set internal limits for reworking.

A single radius cup is the strongest cup configuration and is the most common configuration for round tablets. Adding another radius to the cup changes the cup configuration to a compound or a dual-radius cup. The compound cup has the advantage of having more volume than the single radius cup. Increased volume to the cup will reduce the size of the "belly band," making the tablet appear thinner and easier to swallow. The configuration of compound cup is better for film coating. The rounded edges tend to roll better in the coating pan, reducing the possibilities of edge erosion. There are several disadvantages to using the compound cup design. The intersection of the two cup radii becomes a high-stress point, which is prone to failure under extreme loading, and therefore has a much lower maximum compression force rating than the single radius shallow and standard cup. Extreme loading is not uncommon with the compound cup configuration. The compound cup has more volume; therefore as the upper punch cup enters the die, it fills the die with air, and then must be extracted before compression. For this reason, the compound cup commonly requires slower press speeds or higher compression force than a single radius shallow or standard cup. The compound cup sidewall is steep and receives high abrasion as the tablet is compressed, wearing the tip and weakening the cup. The tip land is critical to the punch tip strength and should be checked often for wear. If the land wears thin, it will cause a condition known as "J hook," which is a common cause of capping and laminating. The land is easily refurbished using 400-grit sharpening stones and a large cotton buff wheel. The compound cup design has a smaller window or available space for engraving and printing than the single radius shallow and standard cup.

Three-dimensional cup configurations are common with vitamins and candies. The three-dimensional cup configuration provides raised features on the tablet surface, providing the opportunity to sculpt features and character details.

33.4.3 Undesirable shapes

A tablet shape too close to round may cause a condition known as punch-to-die binding or self-locking. These shapes must be avoided in order to provide maximum tablet output and satisfactory tool life (Figs. 33.16 and 33.17).

The corner radius of a special shape, such as a square and triangle, is critical for maintaining the strength and integrity of the die. A corner radius less than 0.032″ can cause excessive stress and failure as the die is locked into position with the die lock and subjected to the shock of tablet compression.

33.4.4 Tablet identification

There are two basic methods for identifying a tablet—printing and engraving, with the latter being the most common. There are two styles of engraving—embossed and debossed. With debossing, the identification is raised on the cup face and engraved into the tablet, while embossed identification is cut into the cup face and raised on the tablet (Fig. 33.18). These two styles can be used in conjunction with each other.

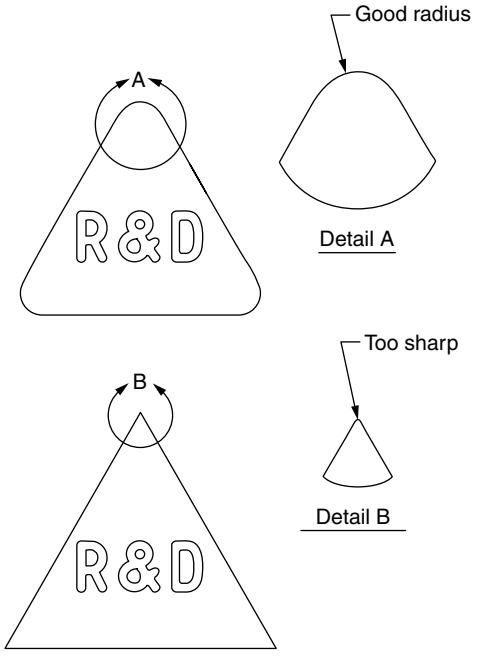


FIGURE 33.16 TSM versus euro method to identify cup radius.

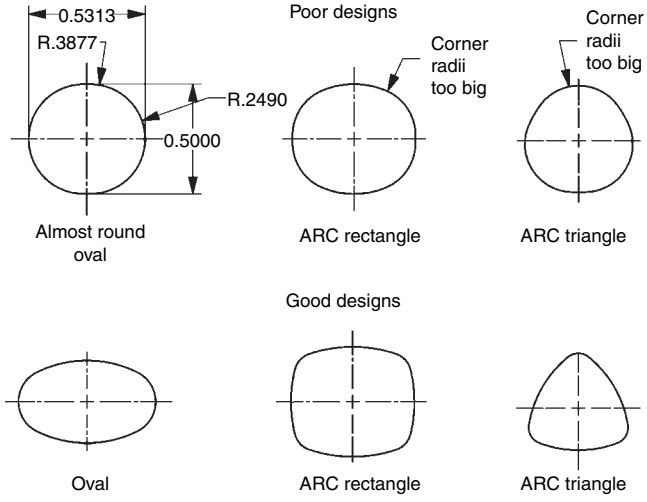


FIGURE 33.17 Good designs versus poor designs.

FIGURE 33.18 Raised embossing in panels.

To ensure product identification, many companies engrave their corporate logos on their product line. As tablet size decreases, the legibility of the identification tends to diminish, eventually reaching the point at which it is no longer legible. For this reason, tablet manufacturers should consider the entire range of tablet sizes when considering the format of a logo for better legibility. As a tablet decreases in size, the logo and drug code are subject to picking (product sticking in or around the identification). Formulation data and product history, if available, should be provided to the tooling manufacturer so that they may engineer an engraving style and format to help minimize picking and sticking because some products are more prone to picking than others.

A company that engraves or embosses most or all of their tablets should consider maintaining a character font. The font should be designed to eliminate sharp corners whenever possible and opening the closed-in areas of a character as much as possible (Fig. 33.19).

For sticky products, the engraving style can be designed to pre-pick the islands of a character, for example, filling in the centers of the "B," "R," "0," "8," etc. The pre-pick character can be difficult to film coat and is prone to fill in and bridging, and therefore the characters can be partially pre-picked for film coated tablets. A partial pre-pick is generally preferred and only removes a percentage of the island instead of removing the island completely. A ramped engraving style, also referred to as a tapered peninsula, provides the same advantage as a pre-picked style and, when used at the outside corners, opens areas of a character. It provides a lower depth of these areas and then tapers the tablet surface (Fig. 33.20).

The radius at the top of an engraving cut at the tablet surface can be a main contributor to picking and tablet erosion. A general guide for the value of the radius is approximately one-third of the engraving cut depth. For example, if the engraving cut depth is .012″, then the radius at the top of the engraving should be 0.003″/0.004″.

FIGURE 33.19 Poor font choice versus good font choice.

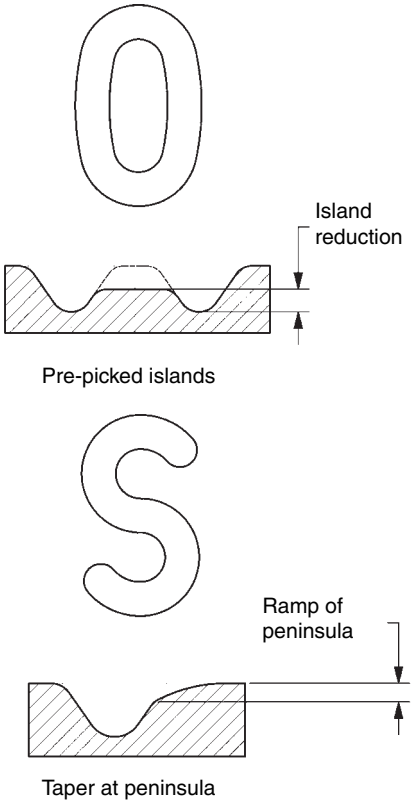


FIGURE 33.20 Partial pre-pick and tapered peninsula.

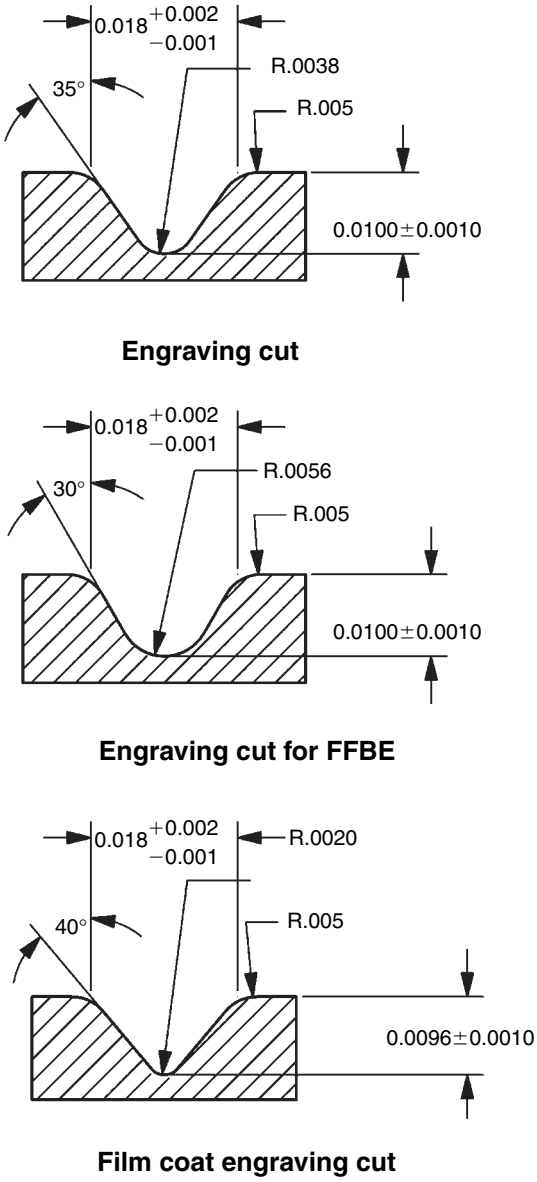


FIGURE 33.21 Engraving cut angles.

The angle of a standard engraving cut for a noncoated tablet is 30 degree. If sticking occurs, it is recommended to increase the angle to 35−40 degree, which is the angle recommended for film-coated tablets. The wider engraving angle may diminish legibility of the engraving cut by allowing more light into the bottom of the cut, but has a better draft angle that provides improved product release (Fig. 33.21).

Incorrectly placing an engraving cut too close to the tablet edge or to the secondary radius for compound cups can result in punch-tip fracturing. Although tooling manufacturers generally maintain certain guidelines for the layout and configuration of the engraving, they must consider the amount of engraving in relation to the tablet size, tablet configuration, and product characteristics before releasing the final tablet design for approval.

33.4.5 Bisects

Bisects, commonly known as a score or break line, are available in a variety of styles.

The purpose of a bisect is to break the tablet into a predetermined dosage, most commonly two equal parts. Breaking a tablet into prescribed dosages should give the consumer a certain degree of confidence that they are receiving the proper dosage. Bisects should be placed on the upper punch whenever possible. Placing the bisect on the lower punch can create problems when the take-off bar removes the tablet from the lower punch. The depth of the bisect is generally deeper than the engraving cut, therefore making it difficult to slide the tablet across the punch face at the ejection cycle.

The standard "TSM" bisect has two different configurations for concave tablets: protruding and cut flush. The protruding bisect style follows the curvature of the cup and extends past the tip edge of the punch. This style helps break the tablet into equal parts because the extended bisect is pressed into the tablet band. The problem with this style is that the protruding bisect may run into the tip edge of the lower punch

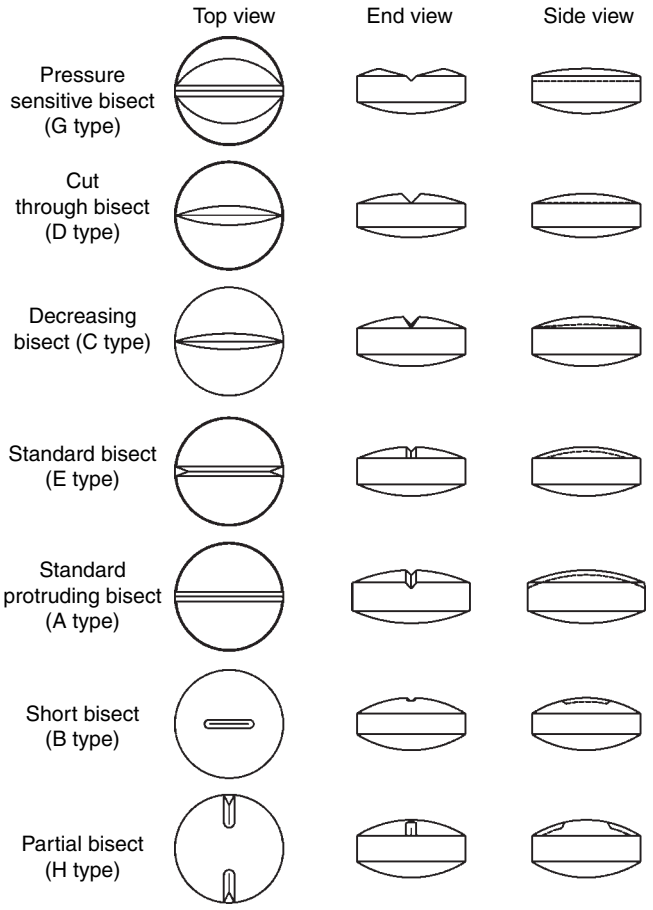


FIGURE 33.22 Bisect types.

if they become too close during tablet press set-up or if the tablet press continues to cycle after the hopper has been emptied. Hitting the bisect into the lower punch edge will leave deep impressions while smashing and swelling the protrusion of the bisect on the upper punch. Standard cut-flush bisect has become more popular for this reason (Fig. 33.22).

A cut-through bisect, also known as a European-style bisect, can only be used on radius cup designs. It has an advantage over the standard bisect by allowing the consumer to easily break the tablet into equal dosages. The cut-through bisect is wider at the center of the tablet than the standard bisect, which reduces the available engraving space on the tablet face. The height of the cut-through bisect is generally the same as the cup depth.

33.4.6 Steel types

Choosing a steel type is generally left up to the tooling manufacturer, unless a specific type has been requested. The criteria for selecting a steel type includes the quantity of tablets to be produced, the abrasiveness or corrosiveness of the granulation, the pressure required for compression, and the cup configuration.

There are two basic categories of steel common in the tablet-tooling industry that are standard and premium. Although the categories may imply that one is superior in quality to the other, this is not the case. Standard steel is the most commonly used grade, whereas the premium grades are used generally for abrasive products and special applications. The cost is generally higher for premium steels due to the quantity of material purchased by the tool manufacturer and the steel composition. Premium steels tend to be harder and more difficult to machine. Premium steels are harder, so they can also be more brittle and prone to fracturing under excessive pressure. Commonly used standard steel grades are: S-5, S-7, S-1, and 408, and premium steel grades D-2, D-3, 440-C stainless steel, and 0–1. Table 33.4 shows the maximum compression.

TABLE 33.4 Maximum Compression

Punch tip diameter	Max compression force by cup depth (kilonewtons)						
	Shallow concave	Standard concave	Deep concave	Extra-deep concave	Modified ball	FFBE	FFRE
1/8	12.5	4.4	2.7	1.8	1.0	3.7	4.9
5/32	18.0	7.0	4.2	3.1	1.6	5.3	7.6
3/16	27.0	9.6	6.1	4.7	2.2	7.2	11.0
7/32	37.0	14.0	8.3	6.7	3.0	9.3	14.9
1/4	49.0	20.0	12.5	10.5	3.9	11.5	19.5
9/32	60.0	27.0	18.5	14.5	5.0	14.0	25.0
5/16	75.0	37.0	26.0	18.0	6.1	16.5	30.0
11/32	92.0	48.0	34.0	22.02	7.4	19.0	37.0
3/8	107.0	61.0	44.0	26.0	8.8	22.0	44.0
13/32	127.2	73.0	55.0	30.0	10.5	25.0	51.0
7/16	149.0	87.0	67.0	35.0	13.5	29.0	60.0
15/32	168.0	104.0	79.0	40.0	14.0	33.0	68.0
1/2	192.0	120.0	92.0	47.0	16.0	38.0	78.0
17/32	219.0	137.0	107.0	53.0	18.0	43.0	88.0
9/16	242.0	159.0	123.0	59.0	20.0	48.0	99.0
19/32	271.0	179.0	139.0	66.0	22.02	53.0	110.0
5/8	302.0	200.0	157.0	73.0	24.0	59.0	122.0
11/16	363.0	246.0	195.0	88.0	30.0	63.0	147.0
3/4	436.0	296.0	238.0	104.0	36.0	75.0	175.0
13/16	509.0	356.0	284.0	122.0	42.0	89.0	206.0
7/8	587.0	417.0	331.0	142.0	48.0	103.0	238.0
15/16	679.0	482.0	286.0	163.0	56.0	118.0	274.0
1	770.0	552.0	445.0	185.0	63.0	119.0	311.0

33.4.7 Inserted dies

Dies are normally manufactured from D-3 premium steel. D-3 steel is superior for wear, but due to its high carbon and high chromium content, does not provide toughness. Dies are not subjected to the same stresses or shock as the punches and, therefore, can be manufactured from harder and better-wearing steels.

The most common die for compressing abrasive granulation is the carbide-lined die. The carbide-lined die has a carbide insert that is inserted into a steel sleeve, which provides a cushion to reduce the possibility of chipping and fracturing the carbide insert. Carbide dies are more expensive than steel dies and are easily justified with the extended die life because die life is easily increased by 10 times in most cases. Carbide die is more brittle and subject to fracturing under excessively heavy compression forces because it is much harder.

If the carbide insert is too thin at its narrowest point due to the tablet size, the carbide insert can fracture due to die-lock pressure and stresses of tablet compression. This is also true for the steel sleeve. The tooling manufacturer should be consulted to determine if a tablet size is acceptable for a carbide liner.

Ceramic-lined dies are becoming more widely used as tougher grades become available. The most common ceramic grade used for compression dies is currently partially stabilized zirconia (PSZ). Dies lined with PSZ have the same general wear characteristics and require the same precautions as carbide-lined dies, but have an advantage in reducing the friction coefficient during the fill and ejection cycles.

When inserting carbide or ceramic dies into the die pocket, a die-driving rod fitted with a nylon tip should be used to prevent carbide fracturing. Die-lock pressure should also be reduced by 10%.

33.4.8 Multitip tooling

Normally, one punch compresses one tablet, with the exception occurring when using multitip tooling. Multitip tools are more common in Europe and have only recently been accepted in the United States. The multitip tool configuration is engineered to compress more than one tablet at a time with the total number of tablets per punch dictated by the punch size and tablet size.

There is a tremendous advantage for using multitip tooling when considering production, operating efficiency, and overall capacity. Increasing production by the multiple of punch tips can be achieved, but typically should not be expected. Using the formula, tablets currently produced $\times$ number of punch tips $\times 0.9 =$ number of tablets expected, will provide a more accurate estimate of the TPM rate.

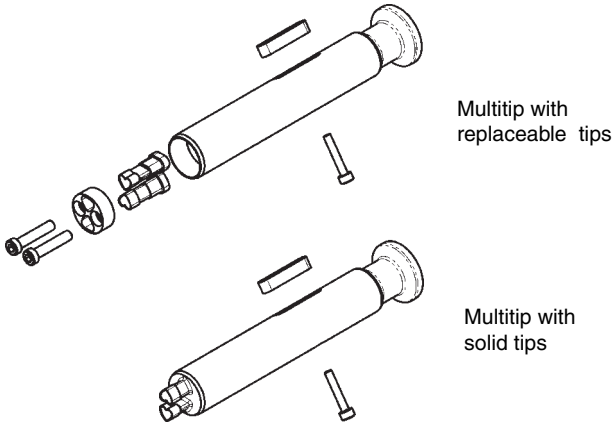


FIGURE 33.23 Multitip tools with replaceable tips and solid tips.

Multitip punches are available in two configurations: a solid punch or an assembly with multiple parts. The solid punch design is easier to clean because it does not have multiple parts, unfortunately, however, if only one tip is damaged the entire punch is unusable and discarded. The multiple part punch design separates the punch tips from the punch body, The punch tips are fixed in place using a cap and/or set screws. If a punch tip is damaged, it can be simply removed and replaced, putting the punch back into service. To properly clean the multiple part design, the tool is required to be disassembled, cleaned, and dried thoroughly before reassembly. This procedure can require substantial labor.

Tablet compression and ejection force most likely will increase, as will the operating temperature. To reduce the potential of product sticking in the punch face, it is suggested to monitor the tablet press and tooling operating temperature. If the compression and/or ejection force becomes too great, the punches may show premature wear on the head flats and back angles and excessive wear on the tablet-press cams.

It is recommended to use the rotating head option for the lower punch. The rotating head will reduce punch in die friction by allowing the punch head to spin freely as it takes the pressure away from the punch tips in the die. The lower punch will be able to move more easily through the cycles of tablet compression (Fig. 33.23).

33.4.9 Punch-tip pressure guide

Punch-tip pressure is calculated by tablet press manufactures and is dictated by the tablet configuration

and steel type. With the assistance of computer-aided designing and finite element analysis (FEA) software, tooling manufactures have become more accurate with the maximum tonnage for round and shaped tablet designs.

Table 33.4 provides the cup configurations with the corresponding maximum tonnage force for round punch tips. This guide has been calculated from the computer-generated procedure "Finite Element Analysis" and is the most accurate guide available.

Calculating the maximum compression force for shaped tablets (ie, capsule oval etc.) can be difficult and confusing. It is recommended to contact the tooling supplier and request these values. The maximum tonnage for round and shaped tablets should be provided on the engineered tablet drawing provided by the tooling supplier, along with the cup volume and surface area. It is important that these values have a strong presence with R&D and are used when formulating a new product. The tonnage requirement should be acceptable before the product reaches the production phase. If tool failure is experienced at the R&D phase, the tablet can be redesigned to accept the required tonnage.

33.5 CARE OF PUNCHES AND DIES

Great care must be taken when cleaning, transporting, and storing punches and dies because they are precision instruments and can damage easily. Upon receiving punches, they should be cleaned and dried thoroughly prior to use. If standard operating procedures require incoming inspection, then the tools should be inspected immediately and any concerns or discrepancies should be reported to the supplier before the tools are used and/or put into storage for future use. Following inspection, the tooling should be lightly oiled, packed in a protective container, and stored in a dry place.

When tooling is required to be shipped, they should not be shipped in storage containers. Most storage containers are not designed to support the weight of the tooling through the handling practices of commercial freight companies. Tooling should be returned in their original individual plastic or cardboard shipping containers and packed tightly to avoid movement. Punch tips are extremely fragile and should be protected at all times from hitting each other or other hard surfaces. A dent or nick on a punch tip can keep the punch from fitting properly into the die. To avoid damage to the die during tablet press set-up, a proper driving rod should be used when inserting the die in the die table. A mild steel rod with the same diameter as the punch guide fitted with a nylon tip is recommended. To prevent damage to the die, die table, and die lock, the die lock pressures indicated by the tablet press manufacturer's operating manual should be observed. Excessive die-lock pressure can distort the die bore and cause punch tightness, fracture the die, and even crack the die table, which would cost thousands of dollars to repair.

33.6 TOOLING INSPECTION

Tooling inspection programs are becoming more common and are performed as a precautionary measure to reassure critical dimensions and embossing details. Confirming critical dimensions will also confirm clearances between the punch and mating parts of the tablet press. Most tooling suppliers will provide a detailed inspection report or a certificate of conformance to assure tablet manufacturers that a specific set of tooling is within the specified tolerance and is capable of producing consistent and quality tablets. The inspection area should be a atmospheric-controlled environment, well-lit for visual inspection, and equipped with calibrated measuring equipment instruments and gauges.

The tooling inspection program should be divided into two sections: incoming inspection and in-process inspection.

The incoming inspection program is for new tools and confirms adherence of critical dimensions. Tools that are supplied with a detailed inspection report should be verified by checking a small percentage of tooling to qualify the supplier's inspection report. A confirmation of the checked dimensions should be recorded and maintained for future reference.

The in-process inspection procedures are recommended for determining wear subjected on critical dimensions responsible for tablet quality and press operation. The most important dimension affecting tablet hardness, weight, and thickness consistency is the working length of the punches. It is not critical to inspect the working length for a calculated dimension, but to instead inspect for consistency within the set. During the inspection process, it is good practice to determine if the punches and dies are in need of polishing and/or light reworking.

The outer dimensions of the punch tip are also critical for inspection and examination. Unfortunately, the worn punch tip is difficult or nearly impossible to inspect using traditional measuring instruments, such as a micrometer or an indicator. The punch tip wears at the edge of the cup and can only be measured accurately using an optical comparator. Dies should be visually checked for wear rings in the compression

zone, and replaced if worn. The severity of a die-wear ring can be checked with an expanding indicator. The expanding indicator will not provide the actual die size, only the depth of the wear ring. The expanding indicator is also capable of measuring the amount and depth of the die taper.

The results of the working length inspection should be documented in addition to noting tool wear and polishing or reworking, if performed. When tooling wear exceeds the new tool specification, it is not generally considered unusable or out of new punch specification.

33.7 TOOLING REWORKING

If considerable reconditioning of the punches and dies is necessary, the tools should be returned to the manufacturer for evaluation. Extensive reworking of the tooling should be performed only by skilled personnel to assure conformance to strict tolerances, providing tablet consistency and proper press operation.

Polishing the cup is the most common procedure of punch maintenance performed by the tablet manufacturer and can be easily performed with proper training. Improper or excessive polishing can reduce the cup depth and diminish the height of the embossing, thus reducing legibility and the ability to film coat. There are three common procedures of polishing the cup: (1) large soft cotton wheel fitted to a bench grinder motor; (2) a dremel tool using nylon brushes and polishing paste; and (3) drag finishing, a process that drags the punch through a fine media of walnut shells or plastic pellets infused with polishing compound. The most effective methods is using the large cotton wheel. Polishing the cup with a large soft cotton wheel is the only method that polishes the cup and restores the critical land at the same time. Restoring the land can increase tool life, strengthen the punch tip, and reduce the likelihood of capping and laminating. Polishing the punch cups with nylon brushes or using a drag finisher is the simplest method of polishing, but will not restore the tip edge or refurbish the tip land, which removes the hooked edge commonly referred to as a "J hook" that is common in capping and laminating. It is not advised to polish or restore the punch head flat because this can alter the critical working length, resulting in inconstant tablet hardness, thickness, and/or weight.

33.8 PRESS WEAR

Tablet press wear can sometimes be the reason for tooling failure and is often overlooked. As the tolerances of punches and dies are constantly monitored, so should the critical tolerances of a tablet press. For example, if tablet overall thickness is inconsistent, the working length of the punches should be checked first; in most cases this dimension is the easiest to check. If the working length of the punches is acceptable, the tools are usually put back into service to frequently experience a reoccurrence of the initial problem. If the pressure roller is out of round, out of concentricity, or worn with severe pitting or flat spots, the result will be inconsistent tablet thickness, which would be expected with improper punch working lengths.

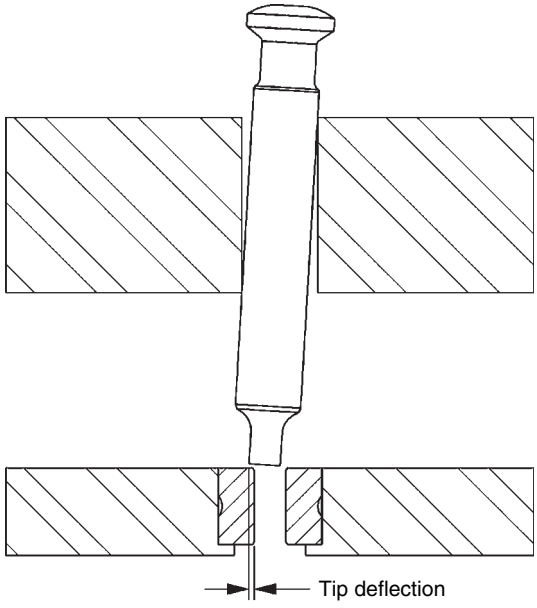


FIGURE 33.24 Deflection of punches indicates excessive wear in turret guides.

Fig. 33.24 shows the correct way to check the turret guide for wear. A new turret may have approximately 0.003″ tip deflection. A turret guide considered worn has a tip deflection of 0.012−0.014″ and should be sleeved or replaced.

Problems in tableting often have a domino effect. It is important to identify and remedy a problem before it affects other areas of the press, the tooling, and the tablet quality.

33.9 PURCHASING TABLET COMPRESSION TOOLING

Considering all the available tooling options and steel types, it can be confusing and complicated to purchase tablet tooling. Before making the final purchase, it is advisable to request a tool and tablet drawing for

approval. Most tooling manufactures can submit a sample tablet made of copper or a special grade of plastic for further approval. The following list is a guideline for information that is needed by the tooling manufacture to properly custom-manufacture tooling that meets the requirements of the product and tablet press:

- The size, shape, and cup depth of the tablet to be compressed (a sample tablet or sample tools would be sufficient if this information is not available)
- Drawing number of the tablet if a drawing exists, if not, request a drawing for future reference
- Hob number, if the order is a replacement
- Press type, model number, and number of stations required
- Steel type if other than standard
- Historical data referencing tablet problems such as capping, sticking, picking, high ejection forces, etc.
- If the tablet is a core and will be coated
- Special options such as tapered dies, domed heads, key type, etc.
- Special shipping instructions

33.10 CONSIDERATION OF TOOLING

Choosing the proper tooling options is critical for a smooth operation. It is recommended to utilize all available industry resources, such as tablet press and tooling manufacturers, for assistance with these choices. It is likely that they have resolved similar difficulties for other customers and have the expertise to recommend the correct options. Recording and maintaining tableting problems are important and should be communicated to the tooling manufacture. Without the knowledge of common tableting problems, the tooling manufacturer can only continue to supply standard tooling.

33.11 APPLICATION OF QUALITY BY DESIGN AND TOOLS (CASE STUDY)

33.11.1 Objective

To develop the compression process parameters and control space for a bi-layer tablet by applying right-first-time (RFT)/six sigma approach.

33.11.2 Methods

The overall approach following the RFT/six sigma is shown in Fig. 33.25. The applied RFT/six sigma tools include failure mode and effects analysis (FMEA), process mapping, critical quality attribute (CQA), critical process parameters (CPPs), cause-and-effect analysis, design of experiment (DOE), control plan, etc.

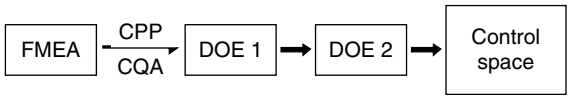


FIGURE 33.25 The overall approach following the RFT/six sigma.

A bi-layer tablet is used for this development. The key ingredient of the formulation includes a water-soluble active pharmaceutical ingredient (API), compressible sugars, flavors, flow-aid, lubricant, and other ingredients. The manufacturing process includes blending, milling and compression using an instrumented rotary press.

FMEA is used to identify CPPs. The CPP's include compression force at first station (precompression), compression force at second station (main compression), and turret speed. The CQAs include potency, content uniformity, and lamination.

Two DOEs are used to understand the knowledge space and identify control space and operation parameters.

33.11.3 DOE design

Study I

Goal: Establish the preliminary compression parameters
Design: Three factors at two levels (3×2)
Factor 1: Main compression
Factor 2: Precompression
Factor 3: Speed

Study II

Goal: Identified optimal process parameters
Design: Two factors at three levels (2×3)
Factor 1: Main Compression
Factor 2: Speed

(Sampling plan: Physician quality reporting initiative-driven)

33.11.4 Results

Lamination is shown in Figs. 33.26 and 33.27.

Based on these results, the operation ranges for compression are optimized as: (1) precompression 0.15–0.3 Kp, (2) main compression 6–7 Kp, and (3) turret speed at 18–20 rpm.

33.11.5 Conclusions

Optimization of compressing of a bi-layer tablet could be difficult due to: (1) the confounding

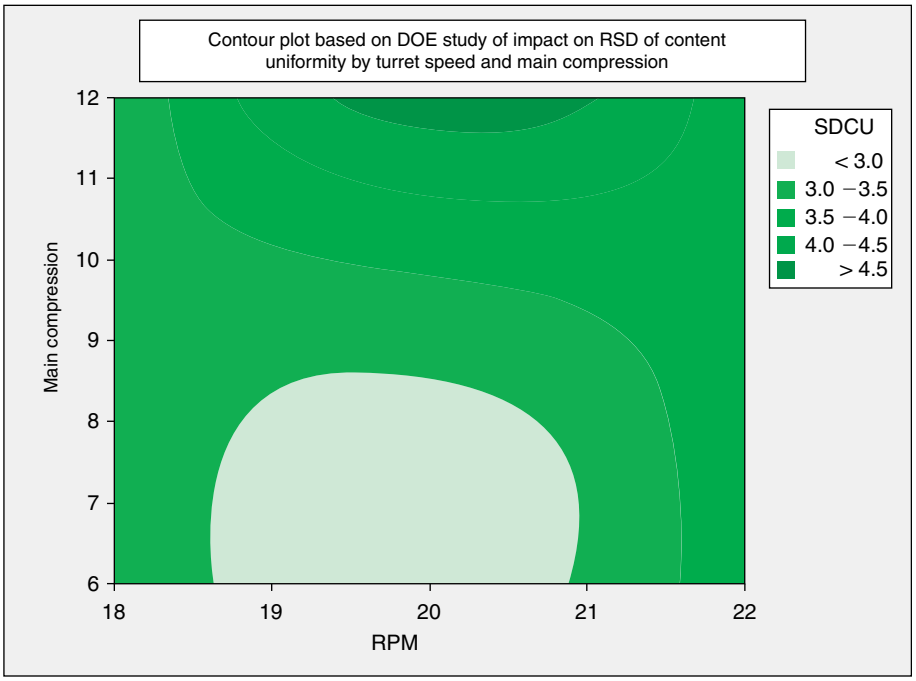


FIGURE 33.26 Impact on RSD of content uniformity by turret speed and main compression force.

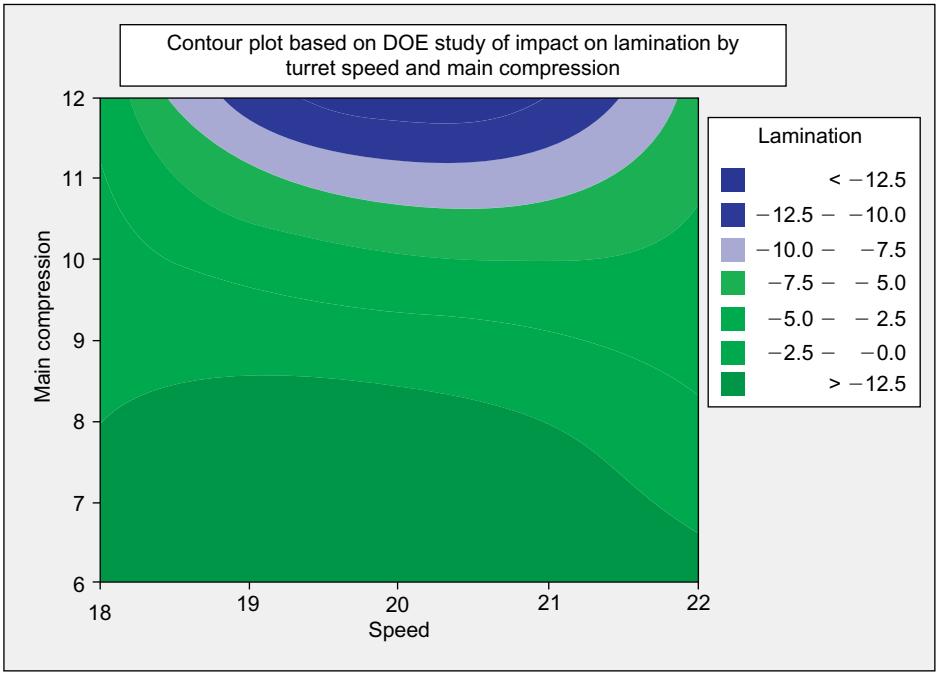


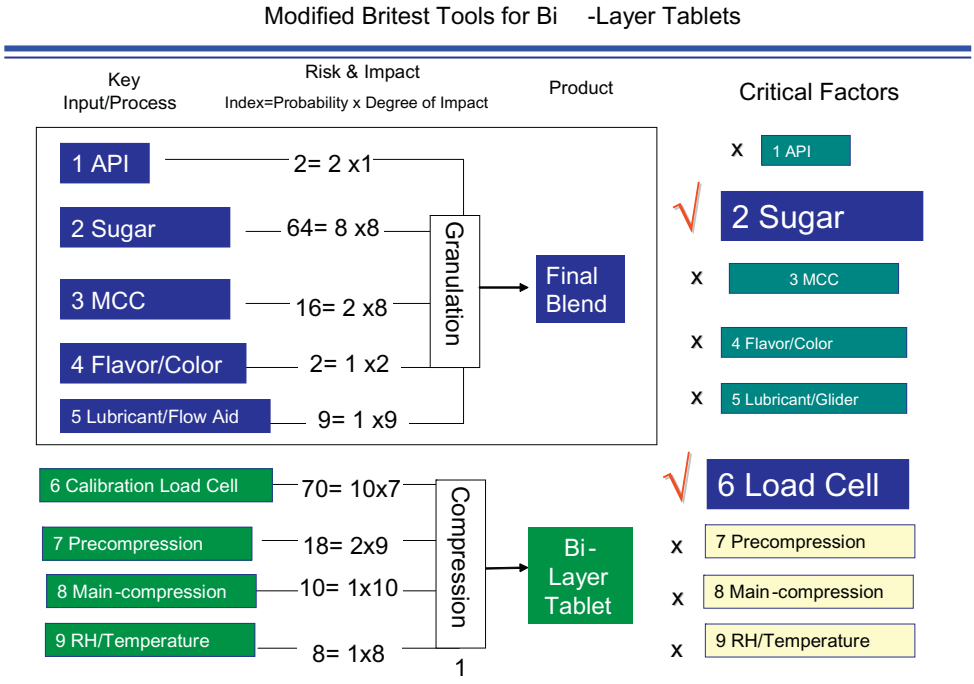
FIGURE 33.27 Impact on lamination by turret speed and main compression force.

Note: Lamination index > 0 = meets requirement

interaction of powder-filling of two layers that may have different flowability, (2) competition of compression force between two layers, and (3) separation/lamination of two layers.

However, by applying the RFT/six sigma approach and utilization of these process characteristic tools, the optimal process parameters and control space can be effectively identified.

33.11.6 Application of britest tool in troubleshooting



33.12 SCALE-UP OF COMPRESSION

In this chapter we will examine press speed as a major factor in tableting, and review measures and functions of tableting speed, such as dwell time, contact time, consolidation time, and decompression (relaxation) time.

The importance of dwell time in compaction has been a subject of discussion for the past 30 years, and yet there is still confusion about how to measure or even define it. Moreover, other important time segments of the compaction event have been mostly neglected because they depend on press geometry and are difficult to quantify.

We have taken a comprehensive approach to calculations of such time events, and found press efficiency factors as by-product. In order to evaluate press or tableting efficiency, we need to consider such important factors as feed time, ejection time, and total tableting time.

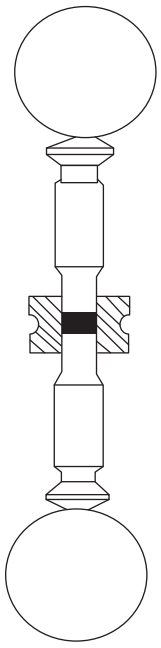


FIGURE 33.28 Punch and die set.

33.12.1 Compaction and compression

Tablets are made of powder compressed in a die by punches. On a rotary tablet press, the die table along with many punches rotates and pushes each set of upper and lower punch between compression rollers (Fig. 33.28). This causes the punches to move inside the die and compress the powder.

Two processes take place when the tablet is made: compaction and compression. During tableting, both processes occur simultaneously (Fig. 33.29).

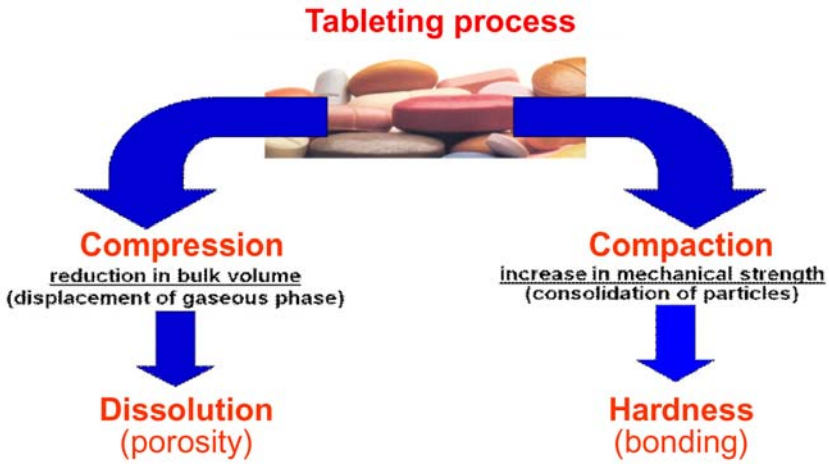


FIGURE 33.29 Compression and compaction as two simultaneous processes in tableting.

By definition, compaction is the increase in mechanical strength of powder under force due to consolidation of particles. Thus, compaction is related to particle consolidation and bonding, which has a direct effect on the tablet hardness and friability.

Compression, on the other hand, is defined as a reduction in bulk volume of the powder under force due to displacement of air between particles. Compression results in a reduction of void space between solid particles, which means a decrease in porosity of a tablet. It is a known fact that porosity, along with the pore size and distribution, affects total surface area, disintegration, and dissolution time.

Several variables are useful in describing the compaction process. Tablet hardness, or breaking force, of a cylindrical tablet can be converted into tensile strength $\sigma = 2 \cdot F/(\pi \cdot d \cdot h)$, where F is the crushing force, d is tablet diameter, and h is tablet thickness (tensile strength is better than hardness because it is normalized with respect to tooling size and shape). A quantity known as "solid fraction", or relative density, is defined as $SF = w/(\rho \cdot v)$, where w is tablet weight, ρ is true density, and v is tablet volume.

Three types of graphs are required to adequately represent and characterize compaction: tabletability (tensile strength vs applied compaction pressure), compressibility (solid fraction vs applied pressure), and compactibility (tensile strength vs solid fraction). If the graphs coincide for any two lots, then formulations are essentially identical with respect to the compaction process. This information can be extremely useful in scale-up.

There are three major mechanisms of particle consolidation: elastic, plastic, and brittle fracture (Table 33.5). Any deformation of elastic materials under stress is temporary; it disappears when the pressure is removed (rubber may provide a typical example). Plastic materials, such as avicel, are deformed permanently even at small pressures. For some materials (such as emcompress), powder particles undergo fragmentation under applied pressure. Such deformation is called "brittle fracture."

TABLE 33.5 Particle Consolidation Mechanisms

Material	Reversible	Time dependent
Elastic (rubber)	Yes	No
Plastic (avicel)	No	Yes
Brittle (emcompress)	No	No
Viscoelastic (starch)	Partly	Yes
Brittle-plastic (lactose)	Partly	Yes

No material deforms by a single mechanism. There is always a combination of either elastic and plastic or brittle and plastic behavior and usually one of the mechanisms predominates. The so-called "viscoelastic" materials deform elastically at low pressures, but as the pressure increases, at some point (yield point) the deformation becomes irreversible. Likewise, there are materials, such as lactose, which initially show some fragmentation, but then may exhibit plastic flow under increased pressure. Plastic flow is time-dependent and decreases as pressure goes up (due to decrease in porosity). Thus, plastic and viscoelastic materials exhibit what is called "strain rate sensitivity," which means they are sensitive to speed of tableting.

As powder particles consolidate, they go through several stages of consolidation (shown in Fig. 33.30 for plastically deforming material). The main events are: reduction in the volume of air between solid particles, particle rearrangement, particle deformation, plastic flow into the interparticulate voids, and formation of interpariculate bonds due to plastic deformation.

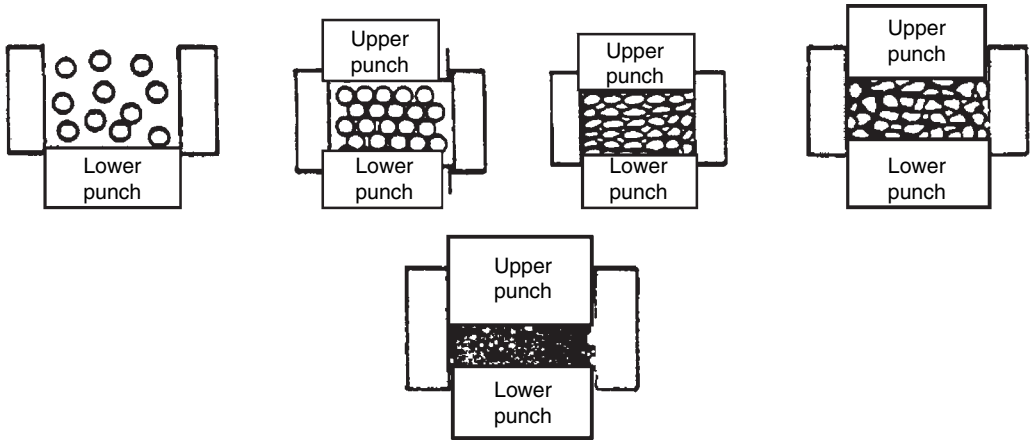


FIGURE 33.30 Consolidation stages with plastic flow.

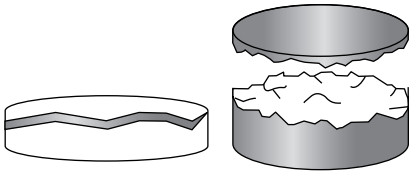


FIGURE 33.31 Tablet capping.

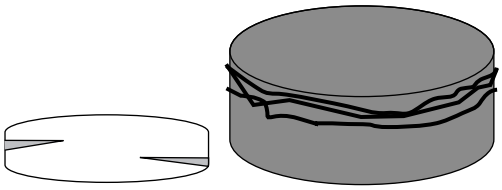


FIGURE 33.32 Lamination.

33.12.2 Tableting failure

Most scale-up problems in tableting are speed-related because the product is moved from a relatively slow R&D press to high-speed production machines. Such problems are: capping, lamination, cracks, picking, and chipping. Other general problems relate to flowability of powders (feeding issues), underlubrication (may cause sticking), and temperature sensitivity (some powders, such as ibuprofen formulations may even melt in the die during long batches as the in-die temperature increases).

Capping is a stratification phenomenon that results in a horizontal dislocation of a tablet layer (Fig. 33.31).

Capping tendency increases with compression force and tableting speed,[1–4] with precompression force,[5] and with punch penetration depth and tablet thickness.[6] Capping occurs due to increase of elastic energy under high-speed compaction compared with a lesser increase of plastic energy. It may also be a result of the expansion of air trapped in pores of the tablet, although this assumption has been disputed in the literature. Press speed is a major factor contributing to capping due to slow process of stress relaxation.

Lamination (Fig. 33.32) occurs when the tablet splits apart in single or multiple layers. Lamination is often blamed on overcompressing—too much compression force flattens out the granules, and they no longer lock together. Lamination can also occur when groups of fine and light particles do not lock together.

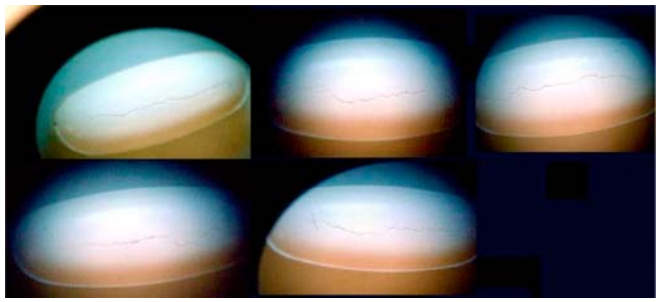


FIGURE 33.33 Crack on the upper side of tablets due to stress relaxation.

Tendency of tablets to laminate increases with speed, compression force, and precompression force.[5,7]

It is a fact that the consolidation part of the compaction cycle (during the "rise time" of the force-time profile) is 6–15 times more important than the decompression part as a factor contributing to capping and lamination.[8,9]

Elastic recovery during decompression can cause stress cracking due to elastic recovery during ejection (Fig. 33.33). This is clearly a strain rate (speed) related phenomenon.

Chipping may be caused by inadequate (brittle) formulation, take-off misalignment, sticking, while picking/sticking to punch faces is formulation-dependent (Fig. 33.34).

33.12.3 Main factors of tableting

The major factors of a tableting process are force and speed of compaction.

The compression force is the dominant factor of the tableting process. As the compression force increases, tensile strength (hardness) of tablets increases, and then may level into plateau or even decrease, as evidenced by numerous studies. Increased force may cause lamination and capping. Friability, disintegration times, and dissolution profile are also affected.

As the punch speed increases, so does the porosity of tablets[10] and their propensity toward capping and lamination. The tensile strength of compacts tends to decrease with faster speeds, especially for plastic and viscoelastic materials, such as starch, lactose, avicel, ibuprofen, or paracetamol; more information can be found in numerous studies.[2,11–20]

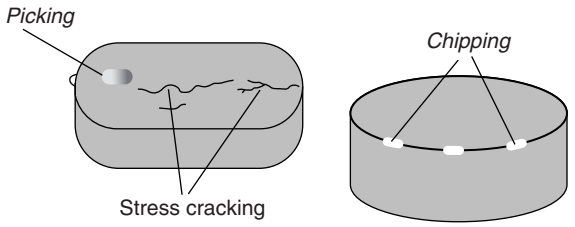


FIGURE 33.34 Picking, chipping, and stress cracking.

Speed also affects the compact temperature in the die, and its mechanical integrity. With increase in porosity, one should expect a drop in disintegration and dissolution times, but the interplay of the force-speed relationship may confound the effect. Although the energy absorbed by the tablet may not change, the power expended in the compaction process may differ greatly with speed, and this, in turn, may have an effect on tablet properties.

The third important factor is force profile, which is directly related to the diameter of compression roll. Larger rolls will provide larger contact times, and with all other factors being equal, this may translate into an increase in tablet hardness (Fig. 33.35 and Table 33.6).

Numerous other factors may affect the scale-up process. Among them is the quality of the measurements, variation in tooling, powder properties, and tablet weight.

TABLE 33.6 Larger Hardness for Larger Roll Diameter at Approximately the Same Dwell Time (Linear Speed)

MATERIAL: AVICEL PH101		
Press:	Betapress	Fette 2090
Stations	16	36
Wheel (mm)	*177.8*	*300*
Speed (RPM)	113	58.2
Dwell time (ms)	9.3	10.2
Hardness (kP)	*24.4*	*36.2*

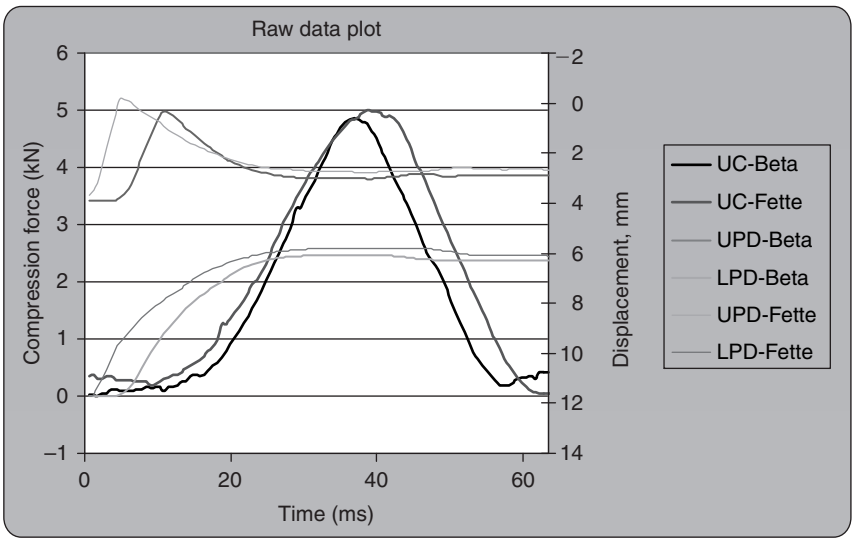


FIGURE 33.35 Traces of upper compression (UC), upper punch displacement (UPD), and lower punch displacement (LPD) for Manesty Betapress and Fette 2090. Material: Avicel.

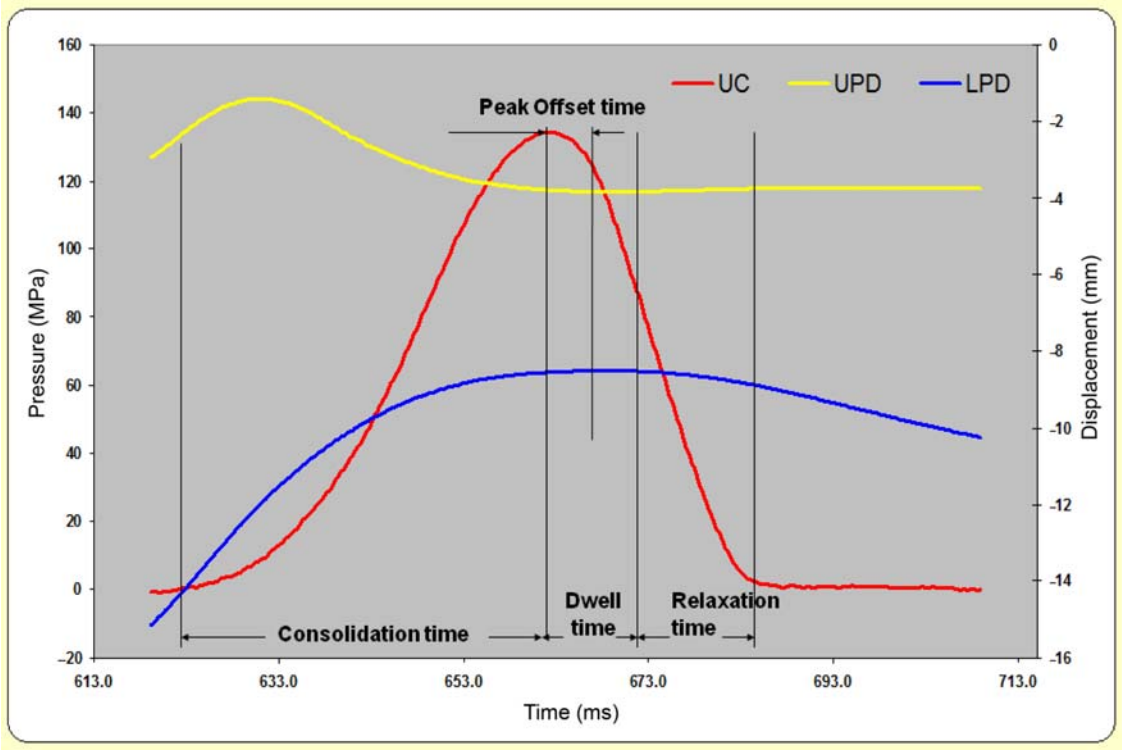


FIGURE 33.36 A compaction event. Legend: *UC*, upper compression; *UPD*, upper punch displacement; *LPD*, lower punch displacement.

33.12.4 Compaction event

A typical compaction event is represented in Fig. 33.36 in terms of pressure (force divided by punch tip area) versus time. The event can be broken into three parts: consolidation, dwell, and relaxation times. We can see that compression reaches its peak well before the punches are vertically aligned with the center of the compression roll (middle of dwell time). The time between the peak of compression and the middle of dwell time is called peak offset time.[21,22] This phenomenon is due to plastic flow in the material being compressed and this flow relieves the pressure, causing the force decrease. Thus, the shape of compression-time form depends on the plasticity of powder.

33.12.5 Tableting time definitions

Let us informally define all discernable time events of the compaction cycle that can be calculated on the basis of geometrical parameters.

- Feeding time, T_f: time when the die is fed with powder
- Contact time, T_c: Time when both punches are moving, having their tips in contact with the material that is being compacted, and their heads are in contact with the pressure rolls.
  $T_c = T_s + T_d + T_r$
- Consolidation (solidification) time, T_s: The portion of contact time, T_c, when punches are changing their vertical position in reference to the rolls, decreasing the distance between the punch tips
- Dwell time, T_d: The portion of Contact Time T_c when punches are not changing their vertical position in reference to the rolls.
- Decompression (relaxation) time, T_r: The portion of contact time, T_c, when punches are changing their vertical position in reference to the rolls, increasing the distance between the punch tips before losing contact with the rolls.
- Compression time, T_p: The portion of contact time, T_c, before the decompression period begins,
  $T_p = T_s + T_d$
- Ejection time, T_e: Time when the tablet is being ejected from the die
- Total time, T_t: time required to produce one tablet on a press (including time between tablets)

Note that all times are defined by punch position relative to the compression roll.

Thus, all calculations can be done on the basis of press geometry, without any reference to product

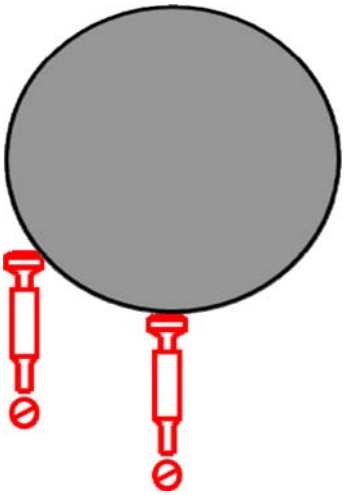


FIGURE 33.37 Geometry of contact time and dwell time.

properties or punch displacement measurements. Once the optimal processing parameters have been established on a research press, the scale-up can be optimized with the help of relatively simple calculations by matching the most important time events.

33.12.6 Dwell time and contact time

As we already know from the stated definition, contact time is the time when the punch head is in contact with the compression roll, while dwell time is when the flat portion of punch head is in touch with the compression roll (Fig. 33.37).

Dwell time is defined as a portion of the contact time when punches are not changing their vertical position with respect to the rolls, that is, when a flat portion of the punch head is in contact with the rolls. Note that dwell time (unlike contact time) as defined does not depend on the roll diameter.

The notion of dwell time is largely misused or misunderstood. In fact, it should be used as a yardstick, a measure of linear (ie, tangential or angular) velocity and therefore will depend on the punch head geometry. The speed comparisons based on the dwell time assume that the punch has a flat head. The velocity then is the length of that flat portion divided by the dwell time. For the same linear velocity, the smaller the punch head flat, the smaller the dwell time. Surprisingly, dwell time for dome-shaped punch heads is practically zero by definition, regardless of press linear velocity or RPM. That is why, compared with dwell time, linear velocity is a better measure of press speed.

Any attempt to calculate dwell time from compression-time traces (eg, as a duration of 90% of the peak) is doomed to failure because such a curve depends on the material properties of the powder being compressed. Plastic flow and elastic recovery distort the "ideal" shape of the force-time profile.

A "classic" simple formula of dwell time does not take into account trajectory curvature:

$$DT(\text{ms}) = (L \cdot NS \cdot 3,600,000) \,/(\pi \cdot PCD \cdot TPH)$$

Where L = Length of a flat portion of the punch head (mm)

NS = Number of stations
π = 3.14159265
PCD = Pitch circle diameter of the turret (mm)
TPH = press speed in terms of tablets per hour

Dwell time is historically used as a sort of "yardstick" to measure compaction speed (Fig. 33.38).

As you can see, Manesty Betapress is ideally positioned within the range of production speeds for the fast-speed presses. That is likely why this press is often used for R&D work. On the other hand, small presses, such as Korsch PH106 or Piccola, do not even come close to benchmark production speeds of 6–15 ms in terms of dwell time.

In fact, dwell time as defined depends on tooling geometry (it is zero for a punch with round head, for example). A linear (tangential) speed of the turret is a better way to represent press speed independently from punch head geometry (Fig. 33.39).

In what follows, we will present the comprehensive set of formulas that represent all tableting events that can be derived from press and tooling geometry.

33.12.7 Tableting geometry

Press speed is a major factor in tableting, and yet the commonly used measures of press speed (such as RPM, tablets per hour, dwell time, linear speed, etc.) do not take into account significant differences in press and tooling geometry, press deformation, tablet thickness, or depth of fill.

Let us introduce a tableting notation that will enable us to generate formulas for exact calculation of the compaction time events (Table 33.7).

In order to visualize some of the variables with respect to press geometry, let us look at Fig. 33.40.

Based on the previous considerations, the following formulas can be derived (Table 33.8).

The previous algorithms and formulas are based on extensive list of published papers and on our own investigations into the geometry and dynamics of tableting.

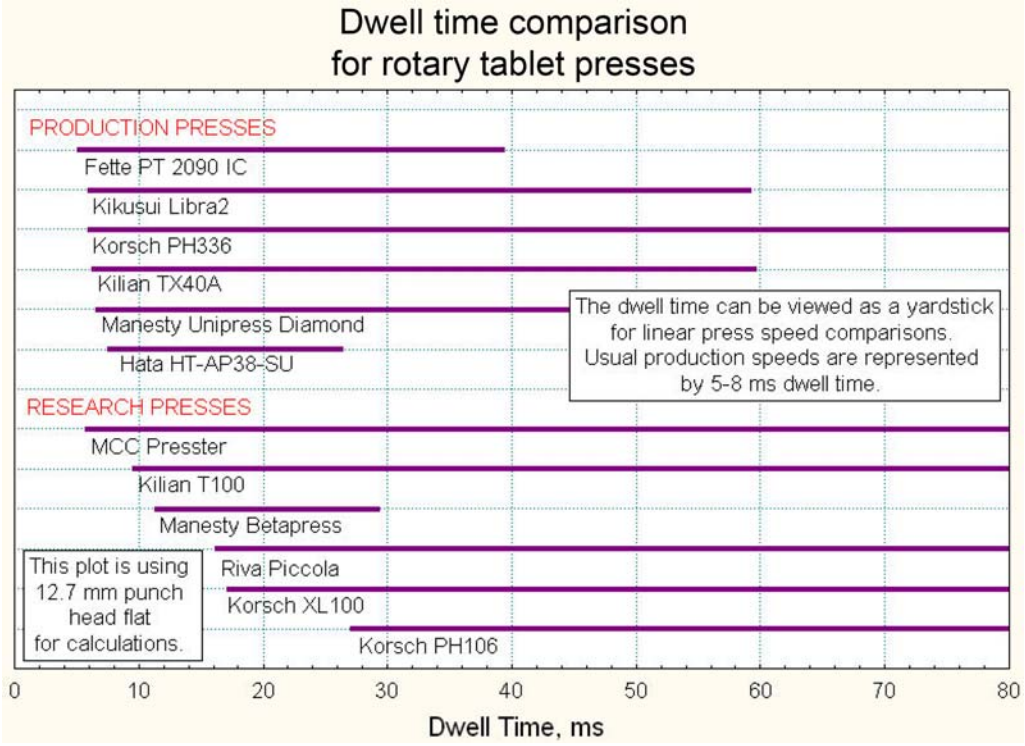


FIGURE 33.38 Dwell time ranges for rotary tablet presses.

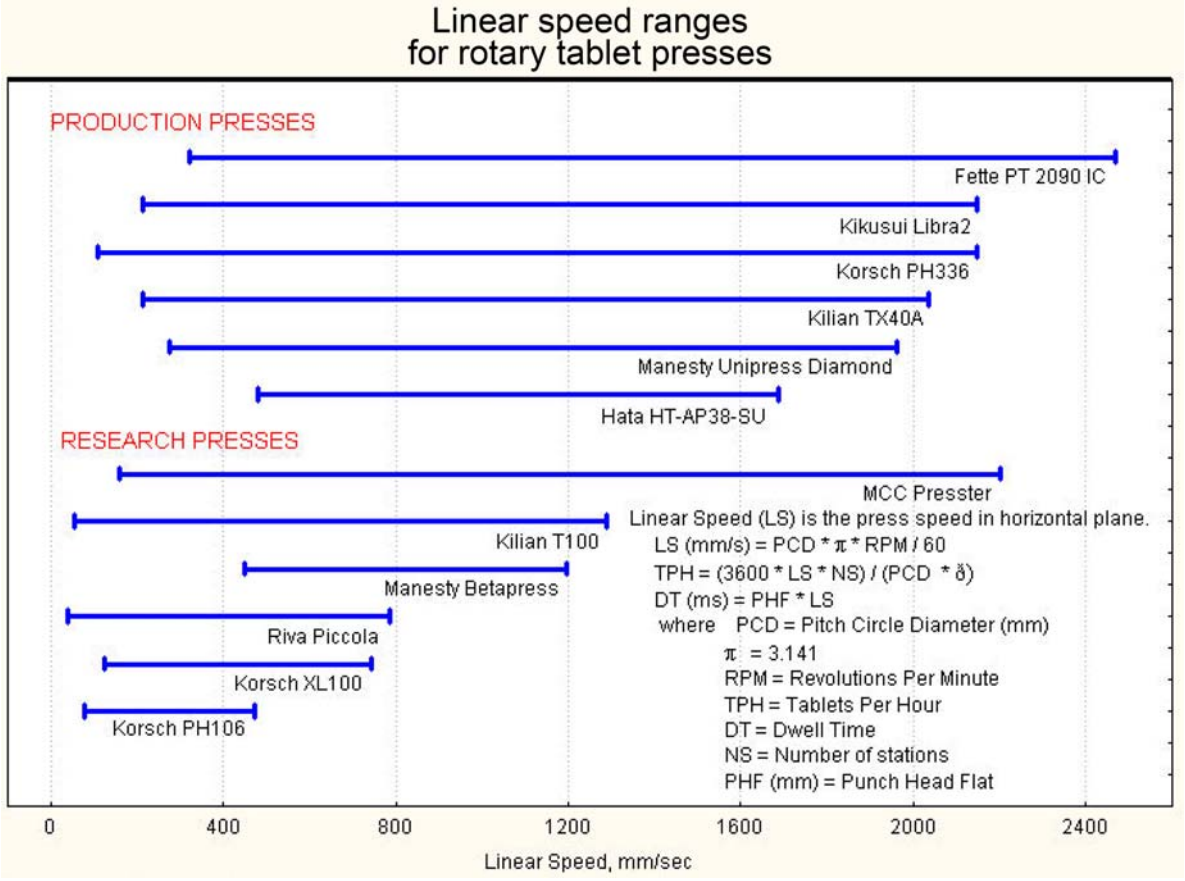


FIGURE 33.39 Linear speed ranges for rotary tablet presses.

TABLE 33.7 Notation Used in Formal Calculations of Compaction Time Events

A	Angular displacement of punch axis from vertical plane through roll axis (rad.)
A_e°	Angle of slope of ejection ramp (deg.)
A_f°	Central angle of turret occupied with feed frame (deg.)
d	Rate of total press deformation (mm/kN)
E_e	Ejection efficiency factor; $E_e = T_e/T_t$
E_f	Feeding efficiency factor; $E_f = T_f/T_t$
E_s	Consolidation efficiency factor; $E_s = T_s/T_t$
F	Compression force (kN)
H_f	Depth of fill (mm)
H_i	In-die tablet thickness (mm)
H_t	Out-of-die tablet thickness (mm)
L_f	Length of feed frame on pitch circle (mm)
N_s	Number of stations
R1	Radius of compression roll (mm)
R2	Radius of punch head curvature (mm)
R3	Radius of punch head flat (mm)
R4	Radius of turret pitch circle (mm)
r	Frequency of turret rotations (1/min)
T1	Time of turret rotation by angle A1 (ms)
T2	time of turret rotation by angle A2 (ms)
T_c	Contact time (ms); $T_c = T_s + T_d + T_r$
T_d	Dwell time (ms)
T_e	Tablet ejection time (ms)
T_f	Die-feeding time (ms)
T_r	Decompression, or relaxation, time (ms)
T_s	Consolidation time (ms)
T_t	Time to produce one tablet on a press (ms)
V_a	Angular velocity of press turret (rad/s)
V_h	Linear horizontal speed of compaction
V_v	Average vertical speed of compaction during consolidation time (compaction rate combined for both punches) (mm/s)
Xs	Horizontal travel of punch during consolidation time (mm) (in direction perpendicular to vertical plane through roll axis)
X1	Horizontal distance of punch axis from vertical plane through roll axis (mm)
Z	Vertical displacement of upper punch from axis of pressure roll (mm)
Zr	Vertical travel of punch during relaxation time (mm); Zr = $[H_t - H_i)/2$
Zs	vertical travel of punch during consolidation time (mm)
Zu	Upper punch penetration (mm)

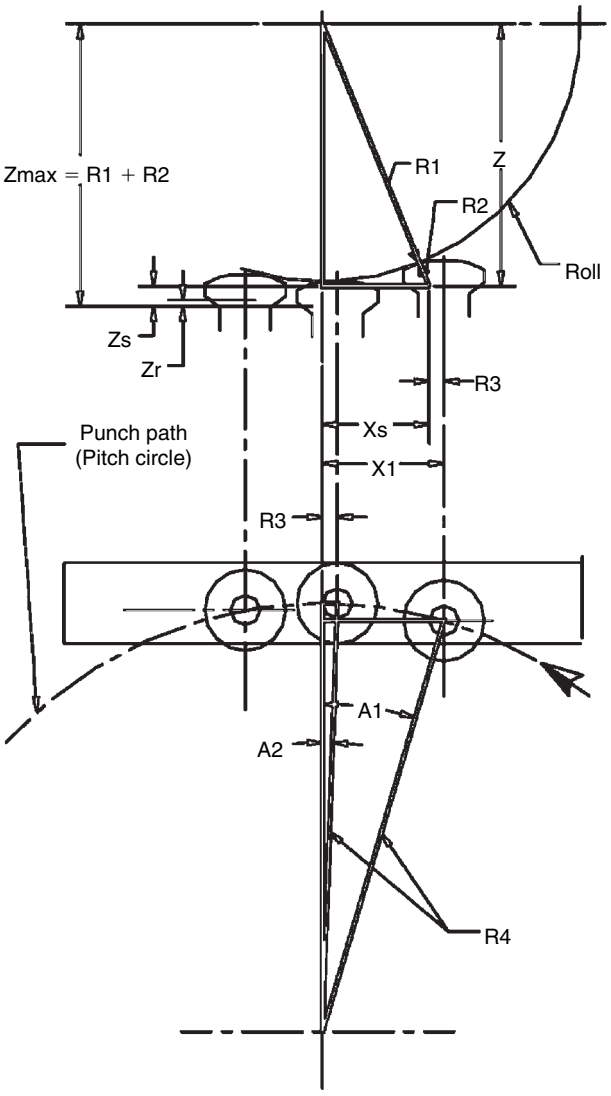
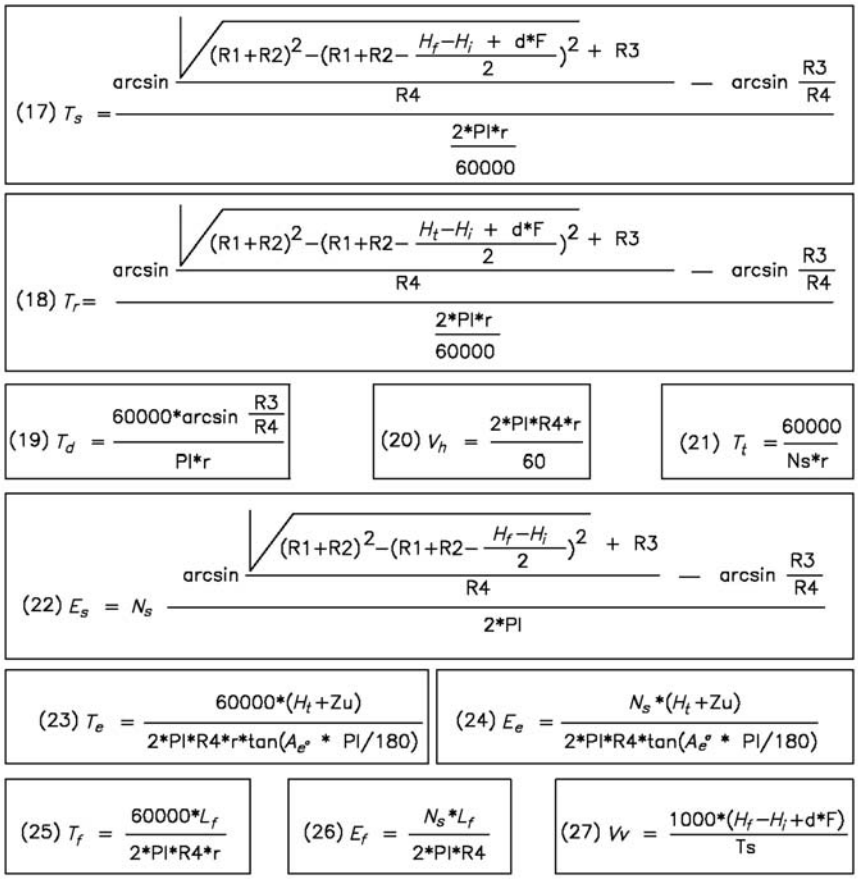


FIGURE 33.40 Graphic representation of variables used in calculation of compaction time events.

Consolidation time, T_s, calculation is based on geometrical parameters such as roll diameter, pitch circle diameter, punch geometry, depth of fill, and in-die tablet thickness. Calculation of T_s takes into account a correction for press deformation if the rate of deformation is known. Other calculation formulas are based on the analysis of Rippie and Danielson[23] and represent an extension of the work by Muñoz-Ruiz et al.[24] Calculation of vertical punch displacement, Zs, during consolidation time is based on depth of fill, H_f, and in-die thickness, H_i. A new, more accurate expression for dwell time, T_d, calculation takes into account the rotational motion of the punches. A proper formula to calculate decompression (relaxation) time, T_r, is used for estimation of all periods of the compaction events and the entire contact time T_c.

TABLE 33.8 Formulas for Tableting times, Velocities, and Press Efficiencies



33.12.8 Tableting scale-up

In order to adequately replicate any process on a different scale, one has to make sure that the processes are similar on both scales. There are three similarity principles in the general theory of modeling: matching geometric, kinematic, and dynamic ratios of characteristic variables. Geometric considerations in simulating compression on a rotary press would require matching die and punch set and compression rolls.

Unlike many other pharmaceutical unit operations, scale-up of tableting process from R&D to production does not involve increase of volume: the die size does not increase. Moreover, it is easy to match the compression force. The most critical factor of tableting scale-up is the rate of compaction (ie, punch speed).

A new class of single-station tablet presses, generally called "compaction simulators," was introduced in 1980s. They were designed to simulate punch movement on a rotary press at high speed to enable making one tablet at a time for formulation development. The classic hydraulic simulator uses a theoretical equation for punch movement, as developed by Rippie and Danielson.[23] Another type of machine, called Presster, is using the principles of similarity to replicate the actual process of making tablets on a rotary press (Fig. 33.41). Yet another product, called Parcus, introduced recently as a material sparing device, can also be used to make single tablets at high production speeds with the goal of eliminating scale-up problems. Useful information on practical use of compaction simulators and their applicability to scale-up can be found in numerous publications (see, eg, Refs. 25–30).

Scale-up of tableting process involves shorter die feeding times (can affect tablet weight), smaller consolidation times (may lower compactibility), and faster ejection (can create cracks or lamination).

At the same production rate, duration of tableting events depends on roll diameter, pitch circle diameter, number of stations, punch geometry, length of feed frame, and the angle of ejection ramp. On the other hand, the speed of various tableting events becomes a limiting factor in press productivity.

However, matching consolidation time T_s, or a sum of T_s and T_d, will, most probably, ensure the same tablet quality, or, at least, bring the process to a close proximity of the target.

FIGURE 33.41 Presster—rotary tablet press replicator. *Source: Photo courtesy of MCC—Metropolitan Computing Corporation, East Hanover, NJ.*

To properly scale-up a formulation based on consolidation time, one should:

1. Evaluate the minimum consolidation time, T_s (or $T_s + T_d$), at which tablets of a satisfactory quality can be produced (ie, the maximum compaction speed) during formulation development. If possible, use a compaction simulator, Presster or a high-speed rotary press equipped with Parcus to establish optimal T_s, or $T_s + T_d$ to match those in production.
2. Calculate at what speed (RPM) different production presses can offer the established minimum satisfactory consolidation time, and what tablet output can be expected from different presses at that speed.

Now that we know how to calculate time events, we can apply this information to practical problems of scale-up. In the following example, we will try to scale-up "a perfect formulation" by matching consolidation time, T_s, and dwell time, T_d.

Let us say that we have a wet granulation of a low dose of brittle API, mixed with Avicel PH102 and 0.5% Magnesium Stearate. On all presses we will use TSM B 3/8″ round flat tooling, with 10-mm depth of fill and 300-mg target tablet weight, out-of-die target thickness of 5 mm, corresponding with 10 kN of compression force. Let us further assume that tablets were made in R&D on a 16-station Manesty Betapress at 50 RPM. Our objective is to move this product to production, where we have a choice of 36-station Korsch PH336, 36-station Kikusui Pegasus 1036, or 37-station Fette P3000 tablet presses.

Calculation of T_s and T_d yields the following results (Table 33.9):

Both the Korsch and Kikusui press can match the $T_s + T_d = 57.6$ ms, but Fette cannot, even at the slowest speed of 30 RPM. If, however, we increase the Betapress speed to 60 RPM, we are able to match $T_s + T_d = 48.1$ for all presses under consideration (Table 33.10). Given the choice, Fette P3000 is preferred because of its high consolidation efficiency factor ($E_s = 1.36$).

Another example involves scale-up of a formulation from 10-station Riva Piccola tabletop R&D press to 14-

TABLE 33.9 Consolidation Time T_s and Dwell Time T_d for Production Tablet Presses Matching Betapress Run at 50 RPM (E_s is the Consolidation Efficiency Factor as Defined Previously)

Tablet press	Stations	RPM	TPH	T_s (ms)	T_d (ms)	$T_s + T_d$ (ms)	E_s
Manesty Betapress	16	50.0	48,000	42.1	15.5	57.6	0.56
Korsch PH336	36	33.4	72,169	44.6	13.0	57.6	0.89
Kikusui Pegasus 1036	36	34.8	75,230	42.6	15.0	57.6	0.89
Fette P3000	37	30.0	133,200	36.7	11.7	48.4	1.36

TABLE 33.10 Matching $T_s + T_d$ for Manesty Betapress, Korsch PH336, Kikusui Pegasus 1036, and Fette 3090

Tablet press	Stations	RPM	TPH	T_s (ms)	T_d (ms)	$T_s + T_d$ (ms)	E_s
Manesty Betapress	16	60.0	57,600	35.1	13.0	48.1	0.56
Korsch PH336	36	40.1	86,603	37.2	10.8	48.0	0.89
Kikusui Pegasus 1036	36	41.8	90,277	35.5	12.5	48.0	0.89
Fette P3000	37	30.2	134,112	36.4	11.6	48.0	1.36

TABLE 33.11 T_s and T_d for Riva Piccola Run at 50 RPM

Tablet press	Stations	RPM	TPH	T_s (ms)	T_d (ms)	$T_s + T_d$ (ms)	E_s
Riva Piccola	10	50.0	30,000	54.1	24.0	78.1	0.80

TABLE 33.12 Matching $T_s + T_d = 69.6$ ms for Riva Compacta, Riva Piccola, Kikusui Libra 2, and Fette PT3090 Rotary Tablet Presses

Tablet press	Stations	RPM	TPH	T_s (ms)	T_d (ms)	$T_s + T_d$ (ms)	E_s
Riva Compacta	18	40.0	43,200	50.0	19.6	69.6	0.60
Riva Piccola	10	56.0	33,612	48.3	21.4	69.7	0.80
Kikusui Libra 2	36	23.9	51,664	51.3	18.3	69.6	0.74
Fette PT3090	61	15.6	114,000	52.7	17.0	69.7	1.67

station Riva Compacta, 36-station Kikusui Libra 2, or 61-station Fette PT3090. Let us assume all formulation, tooling, and processing parameters are the same as in the previous example.

TABLE 33.13 Comparison of $T_s + T_d$ for Fette PT3090 at 60 RPM, Kikusui Libra 2 at 92.2 RPM and Riva Presses at Maximum Speed

Tablet press	Stations	RPM	TPH	T_s (ms)	T_d (ms)	$T_s + T_d$ (ms)	E_s
Fette PT3090	61	60.0	439,400	13.7	4.4	18.1	1.67
Kikusui Libra 2	36	92.2	199,120	13.3	4.8	18.1	0.74
Riva Compacta	18	100.0	108,000	20.0	7.9	27.9	0.60
Riva Piccola	10	100.0	60,000	27.0	12.0	39.0	0.80

For 50 RPM, Riva Piccola achieves $T_s + T_d = 78.1$ ms (Table 33.11). It is too slow for the other three presses to match. Piccola has to be run at least 56 RPM (Table 33.12) to be matched with the other presses at slow speeds (eg, Fette has to be run at a mere 15.6 RPM).

If the production manager and the marketing considerations demand that the product be run on the Fette press (because of its high efficiency and production rate), then we can do the analysis backward to see if the research press can match $T_s + T_d$ of Fette at, say, a mid-range speed of 60 RPM. Table 33.13 shows that while Kikusui Libra can match that target speed, Riva presses cannot.

References

1. Garr JSM, Rubinstein MH. An investigation into the capping of paracetamol at increasing speeds of compression. *Int J Pharm* 1991;**72**:117–22.
2. Garr JSM, Rubinstein MH. The effect of rate of force application on the properties of microcrystalline cellulose and dibasic calcium phosphate mixtures. *Int J Pharm* 1991;**73**:75–80.
3. Mann SC, Roberts RJ, Rowe RC, Hunter BM. The influence of precompression pressure on capping. *J Pharm Pharmacol* 1982;**34**:49p.
4. Mann SC, Roberts RJ, Rowe RC, Hunter BM, Rees JE. The effect of high speed compression at sub-atmospheric pressure on the capping tendency of pharmaceutical tablets. *J Pharm Pharmacol* 1983;**35**:44p.
5. Bateman SD, Rubinstein MH, Thacker HS. Pre—and main compression in tableting. *Pharm Tech Int* 1990;**2**:30–6.
6. Ritter A, Sucker HB. Studies of variables that affect tablet capping. *Pharm Tech* 1980;57–62.
7. Bateman SD, Rubinstein MH, Wright P. The effect of compression speed on the properties of ibuprofen tablets. *J Pharm Pharmacol* 1987;**39**:66p.
8. Mann SC. An investigation of the effect of individual segments of tableting cycle on the capping and lamination of pharmaceutical tablets. *Acta Pharm Seuc* 1987;**24**:54–5.
9. Ruegger CD. *An investigation of the effect of compaction profiles on the tableting properties of pharmaceutical materials*. Ph.D. Thesis, Rutgers University; 1996.
10. Armstrong NA, Palfrey LP. The effect of machine speed on consolidation of four directly compressible tablet diluents. *J Pharm Pharmacol* 1989;**41**:149–51.

11. Fell JT, Newton JM. Effect of particle size and speed of compaction on density changes in tablets of crystalline and spray-dried lactose. *J Pharm Sci* 1971;**60**:1866–9.
12. Rees JE, Rue PJ. Time-dependent deformation of some direct compression excipients. *J Pharm Pharmacol* 1978;**30**:601–7.
13. Armstrong NA, Blundell LP. The effect of machine speed on the compaction of some directly compressible tablet diluents. *Pharmacology*. 1985;**37**:28p.
14. Roberts RJ, Rowe RC. The effect of punch velocity on the compaction of a variety of materials. *J Pharm Pharmacol* 1985;**37**:377–84.
15. Roberts RJ, Rowe RC. The effect of relationship between punch velocity and particle size on the compaction behaviour of materials with varying deformation mechanisms. *J Pharm Pharmacol* 1986;**38**:567–71.
16. Holman LE, Leuenberger H. Effect of compression speed on the relationship between normalized solid fraction and mechanical properties of compacts. *Int J Pharm* 1989;**57**:R1–5.
17. Armstrong NA. Time-dependent factors involved in powder compression and tablet manufacture. *Int J Pharm* 1989;**49**: 1–13.
18. Armstrong NA. Considerations of compression speed in tablet manufacture. *Pharm Tech* 1990;**9**:106–16.
19. Marshall PV, York P, MacLaine JQ. An investigation of the effect of the punch velocity on the compaction properties of ibuprofen. *Powder Technol* 1993;**74**:171–7.
20. Monedero M, Jime Nez-Castellanos MR, Velasco MV, Muñoz-Ruiz A. Effect of compression speed and pressure on the physical characteristics of maltodextrin tablets. *Drug Dev Ind Pharm* 1998;**24**(7):613–21.
21. Dwivedi SK, Oates RJ, Mitchell AG. Peak offset times as an indication of stress relaxation during tableting on a rotary tablet press. *J Pharm Pharmacol* 1991;**43**:673–8.
22. Ho AYK, Jones TM. Punch travel beyond peak force during tablet compression. *J Pharm Pharmacol* 1988;**40**:75p.
23. Rippie EG, Danielson DW. Viscoelastic stress/strain behavior of pharmaceutical tablets: analysis during unloading and post-compression periods. *J Pharm Sci* 1981;**70**:476–82.
24. Muñoz-Ruiz A, et al. Theoretical estimation of dwell and consolidation times in rotary tablet machines. *Drug Dev Ind Pharm* 1992;**18**(9):2011–28.
25. Bateman SD. High speed compression simulators in tableting research. *Pharm J* 1988;**240**:632–3.
26. Bateman SD. A comparative investigation of compression simulators. *Int J Pharm* 1989;**49**:209–12.
27. Celik M, Marshall K. Use of a compaction simulator system in tableting research. Part 1: introduction to and initial experiments with the system. *Drug Dev Ind Pharm* 1989;759–800.
28. Rubinstein MH. Applications of compaction simulators. *Pharm Manuf Int* 1992;177–82.
29. Amidon GE. Powder compaction and "simulation". In: *Arden house conference*, February 308, 2008.
30. Tye CK, Sun C, Amidon GE. Evaluation of the effects of tableting speed on the relationships between compaction pressure, tablet tensile strength, and tablet solid fraction. *J Pharm Sci* 2005;**94**(3):465–72.

This page intentionally left blank

CHAPTER

34

Development, Optimization, and Scale-Up of Process Parameters: Pan Coating

S. Porter[1], *G. Sackett*[2] *and L. Liu*[3]

[1]Ashland Inc., Wilmington, DE, United States [2]Freund-Vector Corporation, Cedar Rapids, IA, United States
[3]Pfizer Inc., Morris Plains, NJ, United States

34.1 INTRODUCTION

In today's pharmaceutical industry, film coating is generally referred to as a process by which a thin, continuous, solid polymer layer is applied to the surface of a dosage form (typically tablets, capsules, and multiparticulates). The purpose of film coating includes esthetic enhancement, increased shelf life, taste masking, moderating the release profile of active pharmaceutical ingredient (API), trade-marking, and protection of intellectual property.

The thickness of the film is generally less than 100 μm. The composition of the film may include a mixture of excipients (polymers, plasticizers, colorants) as well as an API.

34.1.1 The basis of film coating

Film coatings can be applied by different methods, such as spraying a liquid, dipping into a liquid, precipitating from supercritical fluids, or depositing a powder using an electrostatic technique[1].

Spraying a liquid is the most widely used process for film coating and typically includes three basic steps:

1. Spraying an atomized liquid on the target surface that is in continuous movement
2. Maintaining a controlled balance between spray application and evaporation rates by applying heated process air onto the target surface
3. Continuing the process until the desired amount of coating is applied[2]

34.1.2 Evolution of pharmaceutical-coating technology development

Application of coatings to medicinal products can be traced back to the middle ages. However, a uniform and consistent film-coating process that meets the minimum requirements imposed by today's regulatory requirements has only been available for the past few decades.

The earliest process, sugar coating, is performed in a solid wall, smooth bowl-shaped container, employs an external heat source, and is a multistage process that can take several days to complete. Film coating emerged as an alternative to the sugar-coating process to reduce process time.

The evolution in design of coating process equipment has progressed from an open system to one that is completely enclosed and self-contained. Liquid delivery and air-handling systems have also improved as demand for precision and accuracy became more of a regulatory requirement.

The efficiency of heat exchange in film-coating processes has been dramatically improved with the introduction of side-vented (perforated) coating pans that allow the process air to pass completely through the tablet bed, while use of various types of sensors facilitates the monitoring of key process parameters (KPP).

By the end of the 20th century, film coating had evolved to such a degree that it was unrecognizable from processes carried out just 50 years earlier.

Other key developments involve the implementation of process analytical technology, whereby the continual

DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00034-0

assessment of critical process parameters (CPP) that impact critical product quality attributes allows the process to be controlled within a well-defined design space, often obviating the need for routine finished product testing.

Finally, continued process improvements have led to the introduction of innovative coating processes, most notably those involving continuous processing.

34.1.3 Coating equipment introduction

Based on a variety of pan designs, coating pans can typically be classified as solid-wall pans (Fig. 34.1), and fully- or partially-perforated pans (Figs. 34.2 and 34.3).

Based on the type of process, the coating pan can be classified as batch process-coating pan, continuous-coating pan, or off-press continuous coating pan (Fig. 34.4). While early continuous coating pans were designed to coat large batches of tablets (500–2000 kg/hour) where the desire is for large volume throughput at low manufacturing costs, recent designs have focused on slower throughput rates (approximately 50 kg/hour) to match up with the output of a typical production speed tablet press.

FIGURE 34.1 A solid coating pan fitted with immersion sword air-flow enhancement.

34.2 FILM-COATING FORMULATIONS

34.2.1 Overview of types of film-coating formulations

Film-coating systems are usually defined by the way the coating materials are formulated into a liquid coating system, and typically take the form of:

- Polymers applied as organic-solvent-based solutions (typically reserved for modified-release applications today)[3]
- Polymers applied as aqueous solutions (most commonly used for immediate-release film coatings)
- Polymers applied as aqueous dispersions (mainly used for modified-release applications)
- Various materials applied as hot melts (typically used for modified-release and taste-masking applications)

The transition to aqueous processes has been driven by a genuine desire to avoid the hazards associated with using both flammable and potentially toxic solvents, and the added costs of dealing with environmental issues.

Water is not, however, a panacea, and use of aqueous-coating formulations is often associated with challenges that need to be addressed, such as:

- The possibility that processing times will be increased; this issue can, however, be minimized by using high solids aqueous-coating systems[4,5].
- The potentially negative impact on drug stability if water is not effectively removed during processing; however, modern processing techniques coupled with use of higher-solids-coating systems now effectively eliminates this issue.
- The increased likelihood that the harsher process conditions used may affect drug dissolution characteristics; this problem is now minimized by the use of specialized coating formulations that allow processing temperatures to be minimized.

Aqueous-coating formulations can take two forms, namely, aqueous solutions of polymers, which are typically reserved for immediate-release coating applications, and aqueous dispersions (latexes) of polymers, which are frequently used in modified-release coating applications.

When it comes to forming coatings from polymer solutions, the process involves the conversion of a viscous liquid into a visco-elastic solid through a process of continual solvent evaporation, through several stages: (1) Rapid evaporation of solvent

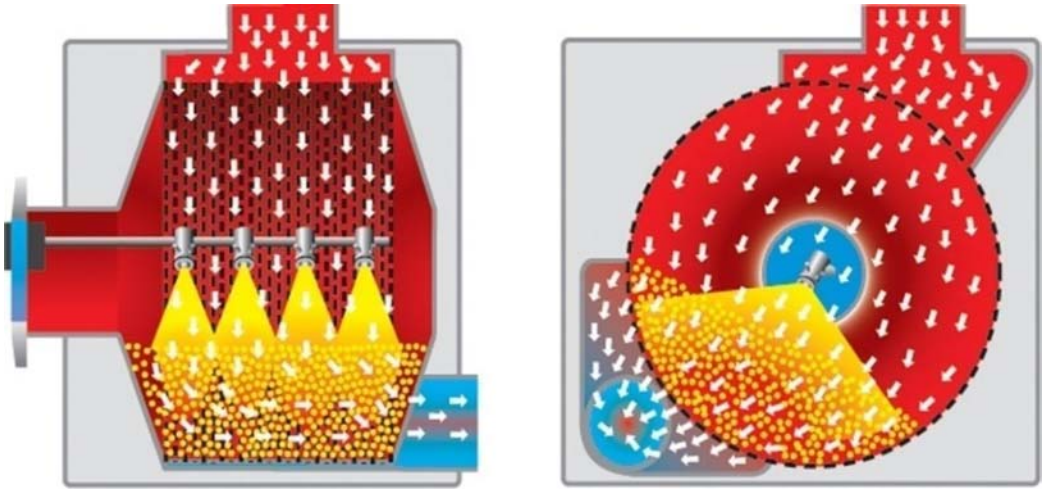


FIGURE 34.2 Schematic of a fully-perforated coating pan. *Source: Courtesy of Vector.*

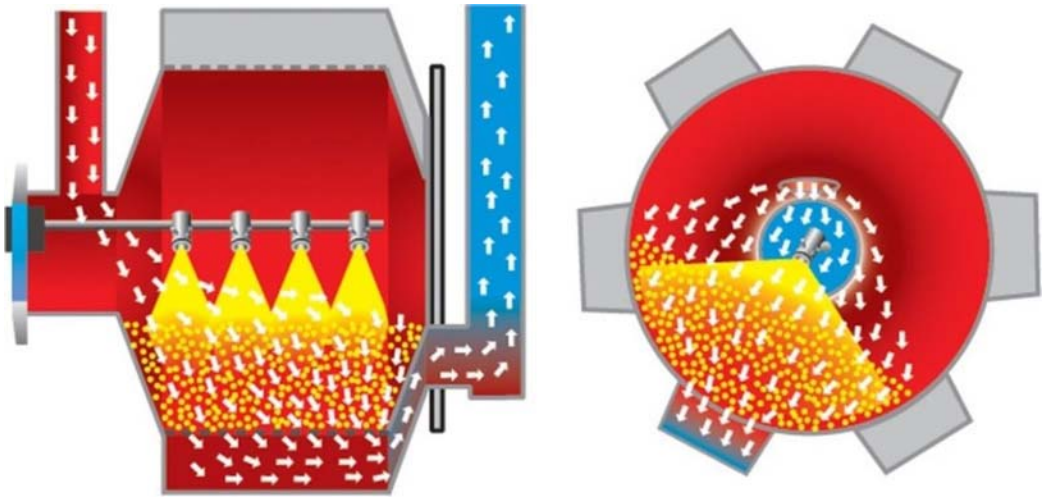


FIGURE 34.3 Schematic of a partially-perforated coating pan. *Source: Courtesy of Vector*[6].

FIGURE 34.4 Example of a continuous coating pan. *Source: Courtesy of O'Hara Inc.*

from finely-atomized droplets of solution, which are deposited onto the surface of the substrate to begin the process of build-up of coating material; (2) Continued solvent evaporation (now at a slower rate) from the coating that is forming on the substrate surface; (3) Immobilization of polymer molecules (at the so-called "solidification point"), such that continued solvent loss now leads to the development of shrinkage stresses within the coating; and (4) Continuation of solvent loss (usually throughout the life of the product) at extremely slow rates.

In contrast, the formation of coatings from aqueous polymer dispersions (Fig. 34.5), while still involving a process of evaporation, is radically different.

Polymer dispersions must undergo a process of "coalescence," where the dispersed polymer particles

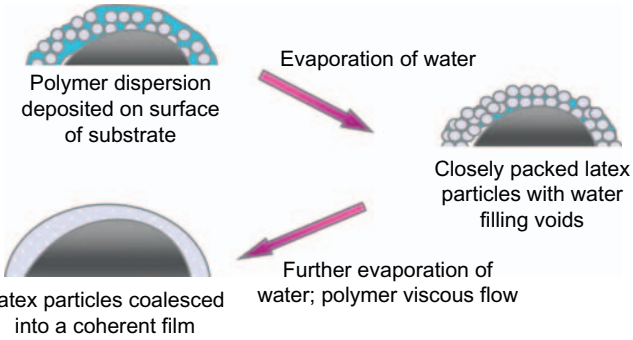


FIGURE 34.5 Schematic of film formation from aqueous polymer dispersions.

must flow together. For this process to occur, pressure develops within the film structure, while the particles of polymer soften under the influence of process heat; as a result, the polymer particles tend to flow together. Pressure development results from evaporation of water; as the porosity of the membrane is reduced during the evaporative process, a capillary network is formed within the structure of the coating, resulting in the development of capillary forces that squeeze the particles together. The ability of the polymer particles to soften sufficiently is dependent on these factors:

- Heat provided by the coating process
- The glass transition temperature of the coating system (dictated by the properties of the polymer, and the properties and concentration of the plasticizer, where needed), which determines the minimum film-forming temperature (MFFT) of the system. The product temperature within the coating pan must exceed the MFFT for film coalescence to occur.

While it is common for many coating formulations to contain a plasticizer (see later discussion), polymer dispersions may require the presence of a plasticizer to facilitate film formation. Generally, it is necessary to use plasticizers with polymer dispersions either when the polymer on which the dispersion is based has a high glass transition temperature (such as ethyl cellulose), so that the MFFT of the system exceeds the product temperatures experienced in the process, or if the flexibility of the resultant film coating needs to be improved.

Unnecessary, or excessive, use of plasticizers should always be avoided, otherwise excessive tackiness will be experienced.

Common ingredients used in many film-coating formulations are: (1) polymer, (2) plasticizer, (3) colorants, and (4) solvent/vehicle.

When designing coating formulations, there are several important issues to consider. These generally involve the need to optimize:

- The visual characteristics of the final product
- The functional characteristics of the coating
- The "processibility" of the coating system (eg, issues relating to preparation and application of the coating liquid, as well as those associated with processing time and costs)

Film coatings are applied to pharmaceutical solid oral dosage forms for a number of purposes, often leading to coatings being categorized on the basis of their intent with regard to influencing drug release, namely immediate-release coatings, which although often used for esthetic purposes, may also be used to improve product stability, facilitate product identification, and achieve an improvement in product organoleptic characteristics (such as taste and odor). Modified-release coatings, which can be subdivided into two sub-categories, namely delayed-release (enteric) coatings, and extended-release (sustained- or controlled-release) coatings.

34.2.2 Overview of types of materials used in film-coating formulations

34.2.2.1 Polymers

Polymers are the essential building block of coating formulations, providing the main characteristics for the final coating formulation[7,8], and are usually characterized in terms of:

- Chemistry, which will mainly influence:
  - Solubility of coating system,
  - Rheology of coating liquid,
  - Mechanical properties of coating, and
  - Permeability characteristics of coating.
- Molecular weight (or molecular weight distribution), which is likely to influence:
  - Mechanical properties of coating and
  - Rheology of coating liquid.

Unlike that of inorganic materials, the molecular weight of a polymer is much more difficult to define because a sample of polymer will consist of a broad distribution of molecular weights.

In order to characterize the molecular weight of a polymer, we usually measure this molecular weight distribution (using techniques such as "gel permeation chromatography"), and then express the molecular weight in terms of one of two statistically defined averages (as shown in Fig. 34.6), namely, weight average molecular weight, M_w, and number average molecular weight, M_n.

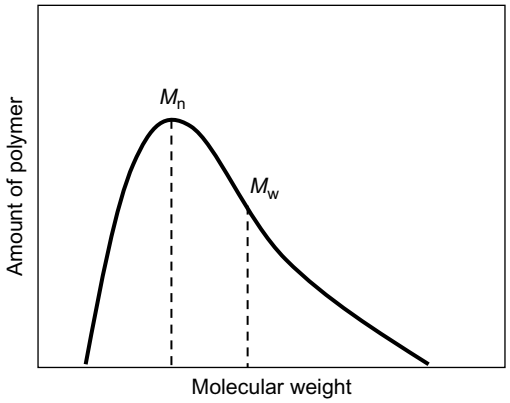


FIGURE 34.6 Typical molecular weight profile for a polymer.

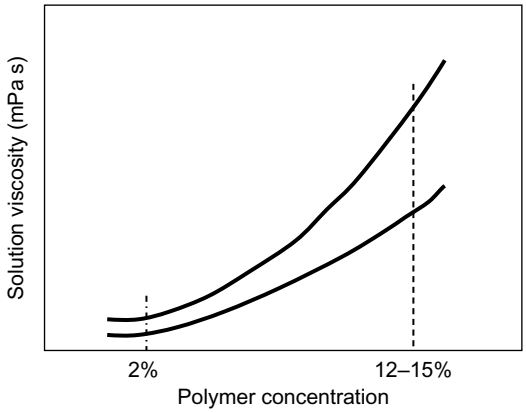


FIGURE 34.8 Effect of polymer concentration on the viscosity of aqueous polymer solutions.

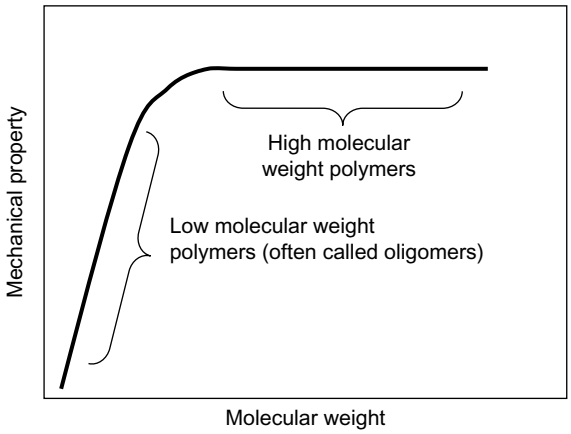


FIGURE 34.7 Effect of molecular weight on the mechanical properties of polymers.

A derived parameter that is often useful is the polydispersity function, which is defined as:

$$\text{Polydispersity} = M_w/M_n$$

The influence of molecular weight on the mechanical properties of polymers is shown in Fig. 34.7, where generally as molecular weight increases, tensile strength and elastic modulus also increase, and adhesion decreases.

Polymer molecular weight can, as stated previously, affect coating solution viscosity. In absolute terms, the relationship between polymer molecular weight and solution viscosity is governed by the Mark–Houwink equation:

$$[\mu] = k\,M^{\alpha}$$

Where $[\mu]$ is the intrinsic viscosity of the polymer, M is its molecular weight, and k and α are constants relating to the solvent system used and the solution temperature.

For more pragmatic reasons, we often prefer to use the term nominal viscosity, where:

$$\text{Molecular weight} = k(\text{nominal viscosity})^{n}$$

Nominal viscosity is usually measured, in water at 25°C, using a solution containing 2% w/w polymer.

Not only does polymer molecular weight influence solution viscosity, but so does polymer concentration, as illustrated in Fig. 34.8. As stated previously, nominal viscosity is measured for dilute solutions of polymer. No particular batch of polymer will conform exactly to the nominal viscosity value for that grade. Thus, the manufacturers (and usually the compendia as well) will set a specification range. As seen in Fig. 34.8, at the 2% solids concentration (usually used for setting the nominal viscosity specification), the batch-wise variation in solution viscosity is usually quite small (often only ±2 cP, or m Pa s). However, as we increase the solids content of the solution, not only does the viscosity increase exponentially, but the difference in solution viscosity from batch-to-batch can increase dramatically, so that the batch-wise variation in coating solution viscosity (depending on the solids contents of our coating solutions) may be as much as 100–200 cP.

34.2.2.2 *Plasticizers*

Plasticizers are another common ingredient added to coating formulations. They are typically used to reduce the *glass transition temperature*, T_g (see Fig. 34.9), and increase coating flexibility.

Most polymers that we use in film coating are essentially amorphous materials, and as such, exhibit a reasonably well-defined glass transition temperature (a fundamental characteristic of polymers that has a profound effect on polymer properties that can also influence film formation, especially when using aqueous polymer dispersions). This transition does not represent a change in state (as we see with melting

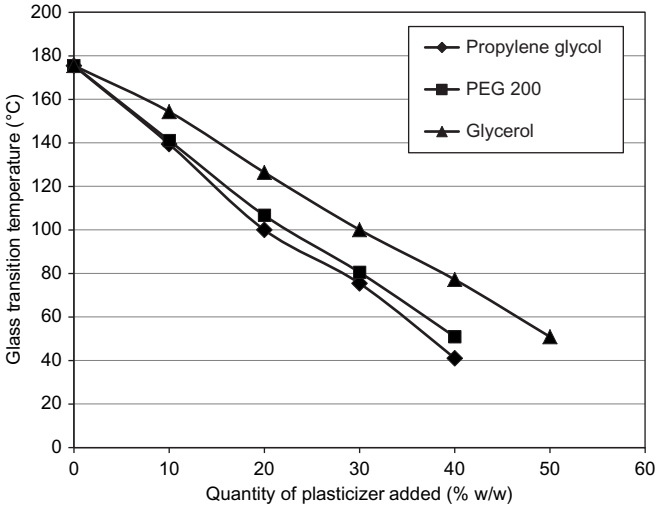


FIGURE 34.9 Effects of plasticizers on the glass transition temperature of HPMC.

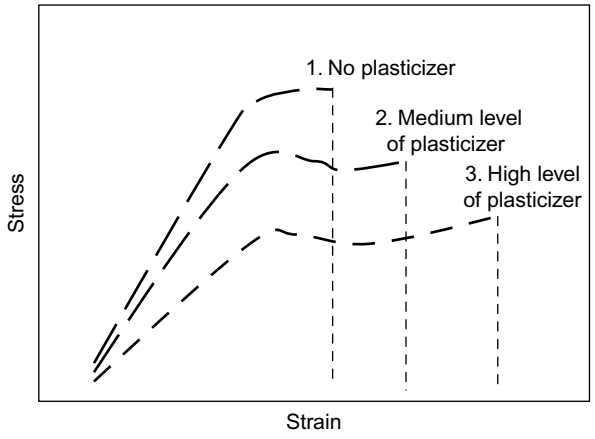


FIGURE 34.10 Effects of plasticizers on the mechanical properties of film coatings.

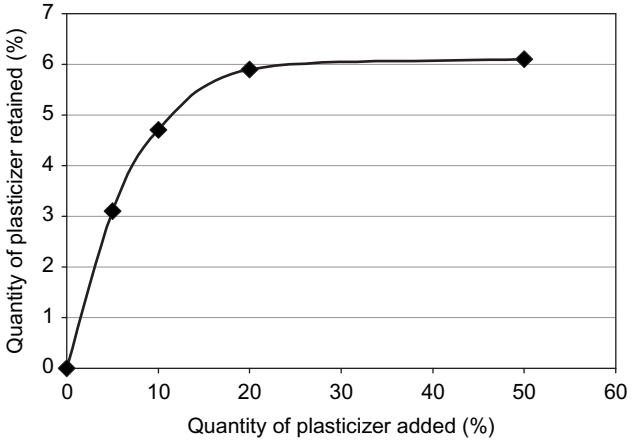


FIGURE 34.11 Lack of plasticizer permanence (for propylene glycol in HPMC films).

TABLE 34.1 Examples of Common Plasticizers Used in Film Coating Formulations

Class	Examples
Polyhydric Alcohols	Propylene glycol. Glycerol. Polyethylene glycols.
Acetate Esters	Glyceryl triacetate (Triacetin). Triethyl citrate. Acetyl triethyl citrate.
Phthalate Esters	Diethyl phthalate
Glycerides	Acetylated monoglycerides.
Oils	Castor oil. Mineral oil.

point), but rather is indicative of a point where there is a dramatic increase in molecular mobility. At this point, the polymer changes from a tough, rigid, inflexible, and brittle material into one that is softer and more pliable. These latter properties are of great value in film coating, and thus it is beneficial to match the glass transition temperature of the final coating system to the coating conditions that will be used.

For most coating systems, it is desirable that the glass transition temperature of that system be optimized for the coating process conditions used. For aqueous polymer dispersions (or latexes), it is critical that such optimization is achieved, otherwise appropriate coalescence of the coating will not occur.

A list of common plasticizers used in coating formulations is shown in Table 34.1, and their comment effects on film mechanical properties are shown in Fig. 34.10.

When it comes to selecting a suitable plasticizer for a coating formulation, some key issues to consider are: efficiency, which defines how much plasticizer must be added to produce the desired effect; compatibility, which indicates how effectively the plasticizer interacts with the polymer, and the level up to which that interaction occurs; and permanence, which relates to both plasticizer-polymer compatibility and plasticizer volatility (see Fig. 34.11 for an example of poor plasticizer permanence).

The common general effects of plasticizers used in film-coating formulations are shown in Table 34.2.

34.2.2.3 *Colorants*

See Table 34.3 for common types of colorant used in film-coating formulations.

Colorants are generally used in film-coating formulations to improve product appearance, aid in product identification, and potentially improve product stability.

While it is possible to use either water-soluble colorants (dyes) or water-insoluble colorants (pigments),

TABLE 34.2 Summary of Common Effects of Plasticizers

Property	Effect of increased plasticizer concentration
Tensile strength	Reduced
Elastic modulus	Reduced
Film adhesion	Variable, but increased under optimal-use conditions
Viscosity of coating liquid	Usually increased, with effect being greater as plasticizer molecular weight is increased
Film permeability	Variable, depending on physicochemical properties of plasticizer
Glass transition temperature (T_g) of film	Reduced, with magnitude of effect being influenced by compatibility with polymer

TABLE 34.3 Common Types of Colorant Used in Film-Coating Formulations

Type	Examples
Water-soluble dyes	• FD&C Yellow #5 • FD&C Blue #2
FD&C lakes	• FD&C Yellow #5 Lake • FD&C Blue #2 Lake
D&C lakes	• D&C Yellow #10 Lake • D&C red #30 Lake
Inorganic pigments	• Titanium dioxide • Iron oxides
"Natural" colorants	• Riboflavin • Beta-carotene • Carmine lake

TABLE 34.4 Contrast Ratio Values for Coatings Colored With Selected Pigments

Pigment	Contrast ratio
None	33.3
Titanium dioxide	91.6
Red iron oxide	99.5
Yellow iron oxide	98.4
Indigo carmine lake	99.5
Tartrazine lake	66.7

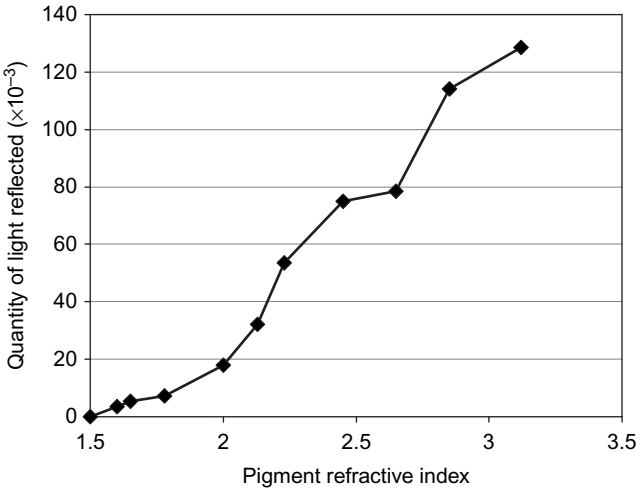


FIGURE 34.12 Hiding power and influence of pigment refractive index.

as shown in Table 34.3, the use of pigments has several advantages because they:

- Replace some of the polymer, thus allowing the solids content of film-coating liquid to be increased while still maintaining a sprayable viscosity;
- Act as a solid inclusion in the final dried coating, often improving the moisture barrier properties of the coating;
- Exhibit better light stability than water-soluble colorants;
- Are less prone to color migration as the coatings dry; and
- Exhibit light excluding behavior, thus improving the stability of photo-labile drug substances

In order to exert these properties effectively, it is desirable that the colored film coating is able to completely mask the substrate. The ability to achieve this goal is usually expressed as the hiding power (or contrast ratio) of the coating. The consequences of poor hiding power include the fact that higher coating levels will have to be employed to achieve uniform appearance from tablet to tablet, and from batch to batch, and if the tablets are intagliated (ie, have a logo), higher levels of coating can cause increased risk of logo bridging.

The contrast ratios of film coatings colored with selected pigments are shown in Table 34.4.

Generally, the main factors that can influence the hiding power of a particular film coating are:

- The quantity of light reflected at the polymer–pigment interface (which, in turn, is influenced by the refractive index of the colorant, see Fig. 34.12)
- The wave length of light absorbed by the colorant
- The amount of light absorbed
- The concentration of the colorant in the coating (see Fig. 34.13)
- The thickness of the coating (see Fig. 34.13)

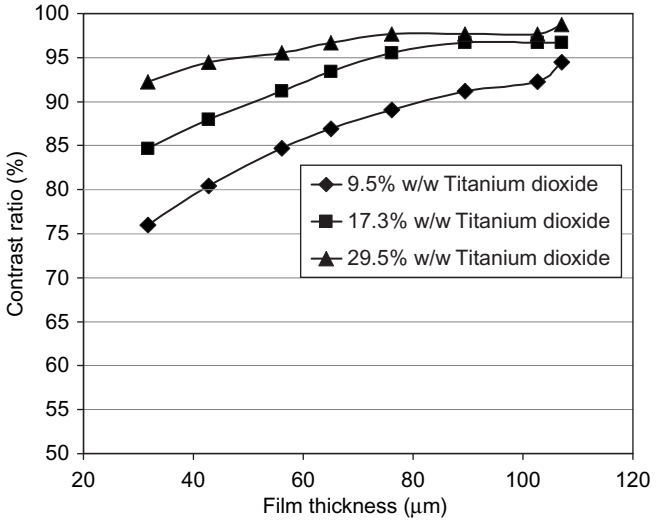


FIGURE 34.13 Effects of pigment concentration and coating thickness on the contrast ratios of colored film coatings.

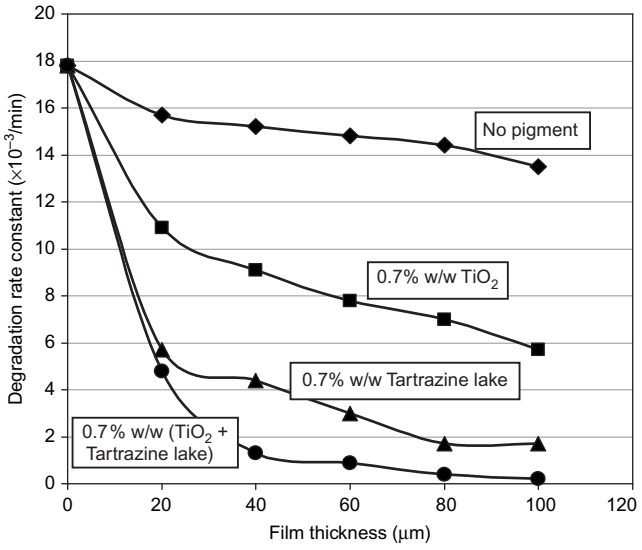


FIGURE 34.14 Photostabilization of nifedipine using pigmented film coatings.

The ability of film coatings to improve product stability is generally related to:

- The influence of pigments on coating permeability (generally, as pigment concentration is increased, up to a critical level called the critical pigment volume concentration, film permeability to environmental gases such as water vapor and/or oxygen are reduced); decreased permeability can lead to improved product stability
- The ability of the coating to exclude light, an important issue when coating products containing photolabile APIs (such as Nifedipine, as shown in Fig. 34.14)

In summary, the general effects of pigments on the properties of film coatings are shown in Table 34.5.

TABLE 34.5 General Effects of Pigments in Film Coatings

Property	Effect of increased pigment concentration in film
Tensile strength	Reduced (but effect may be minimized by effective pigment dispersion in film)
Elastic modulus	Increased
Film adhesion	Generally, little effect
Viscosity of coating liquid	Increased, but usually not substantially
Film permeability	Reduced, unless critical pigment volume concentration (CPVC) is exceeded
Hiding power	Increased, but effect is dependent on refractive index and light absorption characteristics, of pigment

TABLE 34.6 Examples of Common Solvents Used in Film-Coating Formulations

Class	Examples
Water	–
Alcohols	• Methanol • Ethanol • Isopropanol
Esters	• Ethyl acetate • Ethyl lactate
Ketones	• Acetone
Chlorinated hydrocarbons	• Methylene chloride • 1:1:1 Trichloroethane • Chloroform

34.2.2.4 *Other additives*

While polymers, plasticizers, and colorants constitute the major ingredients in film-coating formulations, other materials that might be used for a variety of reasons include:

- Antiadhesive agents, especially when the polymer is somewhat tacky;
- Flavoring agents (typically in nutraceutical applications);
- Surfactants (to facilitate the wetting of insoluble materials dispersed in the coating formulation);
- Pore-forming agents (especially in extended-release film-coating formulations).

34.2.2.5 *Solvents/vehicles*

Currently, most coating formulations that are applied to solid oral dosage forms are liquids. Thus, the key additive that is used to render the coating formulation a liquid is a solvent (the term vehicle is used to describe the situation where an aqueous polymer formulation is preferred).

A list of common solvents used in coating formulations is given in Table 34.6.

Quite clearly, in modern film-coating practices, water is often the preferred solvent[9]. Nonetheless, water is a compromise solvent for these reasons:

- Water has higher latent heat of vaporization (see Table 34.7), thus requiring more energy input to the coating process to assure that effective drying takes place (thus often precluding the use of conventional coating equipment).
- Aqueous coating systems have higher surface tensions than their organic-solvent-based counterparts (see Table 34.8), thus impacting wetting and adhesion (see Table 34.9) on many types of pharmaceutical tablets (eg, vitamins).
- Aqueous coating systems are more viscous than organic solvent-based systems (see Table 34.10), thus having some impact on pumping and atomization efficiency.

TABLE 34.7 Latent Heats of Vaporization of Common Solvents

Solvent	Latent heat of vaporization (kJ/kg)
Methylene chloride	556.7
Methanol	1967.1
Water	2260.4

TABLE 34.8 Surface Tension Values of Common Solvents

Solvent	Surface tension (mN/m) at 20°C
Acetone	23.7
Chloroform	27.1
Ethyl alcohol	22.8
Methyl alcohol	22.6
Water	72.8

TABLE 34.9 Adhesion Values of HPC Coatings Deposited From Various Solvents

Solvent system	Adhesion value (kPa)
Methylene chloride:methanol (9:1)	29.8
Ethanol:water (95:5)	23.9
Acetone:water (9:1)	20.2
Chloroform	14.4
Water	7.1

TABLE 34.10 Influence of Solvents on the Viscosity of HPMC-Based Coating Solutions

Solvent	Polymer concentration (% w/w)	Solution viscosity (m Pa s) (cP)
60:40 methylene chloride–methanol	5.0	40
80:20 ethanol–water	5.0	70
Water	10.0	450

In some cases, use of organic solvents has continued, especially when:

- The coating process will not accommodate the use of water (ie, drying is poor);
- The adhesion achieved with aqueous systems is unacceptable;
- Certain critical ingredients (eg, polymer) are neither water-soluble nor available as a latex system; and
- Exposure to an aqueous process would cause stability problems for the product being coated.

34.2.3 Film-coating formulations used for immediate-release applications

34.2.3.1 Characteristics of polymers used

- Generally water soluble;
- Able to form strong, flexible films;
- Adhere strongly to tablet surfaces;
- Form elegant films;
- Facilitate ease of processing (pumping, spraying, atomization, and lack of tackiness); and
- Permit rapid drug release from dosage form.

A typical dissolution profile for a product coated with an immediate-release film coating is shown in Fig. 34.15.

34.2.3.2 Examples of types of polymers used

Common polymers used in immediate-release coating formulations are listed in Table 34.11.

34.2.3.2.1 Cellulosic polymers

Cellulosic polymers, especially hydroxypropylmethylcellulose (HPMC), have long been the mainstay of film-coating formulations, with a popularity that stems from their common usage from the early days of film coating when organic solvents were always used; global regulatory acceptance; ready availability from a number of vendors; and ability to form coatings generally having acceptable properties (such as good film strength and aqueous solubility).

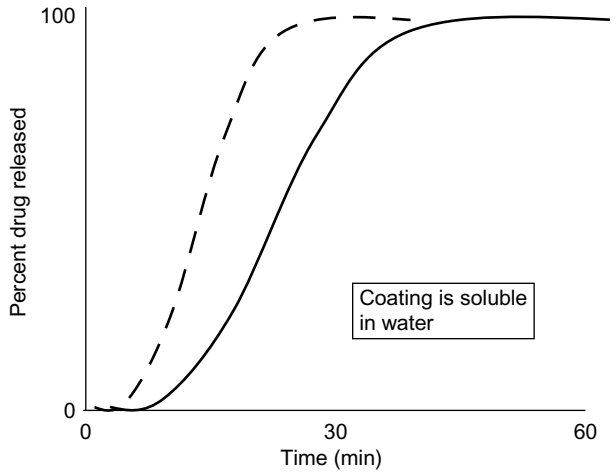


FIGURE 34.15 Typical drug release profile for products coated with immediate-release coatings.

TABLE 34.11 Examples of Common Polymers Used in Immediate-Release Coating Formulations

Polymer class	Examples
Cellulosic	• Hydroxypropylmethylcellulose • Hydroxypropylcellulose • Hydroxyethylcellulose
Vinyl	• Poly (vinyl pyrrolidone) • Poly (vinyl alcohol) • Poly (vinyl pyrrolidone)−poly (vinyl acetate) copolymers • Poly (vinyl alcohol)−poly (ethylene glycol) copolymers
Glycols	• Poly (ethylene glycol)
Acrylics	• Amino alkyl methacrylate copolymers
Other carbohydrates	• Maltodextrin • Polydextrose

34.2.3.2.2 Vinyl pyrrolidone polymers

The most common vinyl polymer used in the pharmaceutical industry today is poly (vinyl pyrrolidone). While this polymer has been primarily used as a wet binder in granulation processes, it has certain utility in film coatings (due to its potentially high film-adhesion characteristics), although uses are somewhat limited because it generally forms extremely tacky films (both during application of the coating and on final coated product), and produces coatings that tend to be somewhat brittle.

In contrast, poly (vinyl pyrrolidone)-poly (vinyl acetate) copolymer potentially has greater utility because it is less tacky than the homopolymer, produces films that exhibit good adhesion characteristics, and generates coating solutions with low viscosities.

34.2.3.2.3 Vinyl alcohol polymers

Poly (vinyl alcohol) has gained popularity recently because of its good film properties, and the relatively low viscosity of its coating solutions. Typical characteristics of coatings made with this polymer include the fact that coating solutions that can be somewhat tacky, while subsequent dry films tend to exhibit good adhesion properties (to tablet surfaces) and excellent barrier properties (to environmental gases such as oxygen and water vapor); however, applied coatings may, under certain circumstances, slightly retard tablet dissolution behavior.

To obviate some of the problems of the homopolymer (especially tackiness), poly (vinyl alcohol)-poly (ethylene glycol) is a copolymer designed with certain improvements in mind.

34.2.3.2.4 Acrylic polymers

The acrylic polymers that are typically used in film-coating applications are:

- Generally not water soluble, per se (the polymer tends to dissolve readily at low pH);
- Traditionally used to create film coatings with improved taste-masking capabilities because the pH in the mouth is usually alkaline (ie, above that where the coating will dissolve in water); and
- Usually applied as solutions in organic solvents, although special polymer grades that allow aqueous polymer suspensions to be prepared are now available.

34.2.3.3 Formulation strategies used

Immediate-release film coatings have traditionally utilized relatively simple formulation strategies, typically combining a single polymer in combination with other ingredients, such as plasticizers and colorants. The introduction of aqueous coatings created a major challenge for formulators, especially with regard to managing the viscosity of the coating formulation and achieving acceptable adhesion of the coating to the surface of the tablets being coated.

Such challenges have typically been met, over the course of time, by utilizing polymer mixtures, such as blends of:

- Different molecular weight grades of the same polymer (such as a mixture of HPMC 6 cP and HPMC 3 cP),
- Similar polymers (such as HPMC + HPC), and
- Cellulosic polymers with other carbohydrate materials (such as HPMC with maltodextrin, polydextrose, or lactose).

The properties of some typical polymer blends are shown in Table 34.12.

TABLE 34.12 Film Properties of Various Polymer Blends

Polymer	Tensile strength, σ (MPa)	Elastic modulus, E (GPa)	σ/E value ($\times 10^{-2}$)	Film adhesion (kPa)
HPMC blend	31.74	2.25	1.41	250.5
HPMC/HPC blend	17.68	0.99	1.79	253.5
HPMC/lactose blend	13.53	2.34	0.58	500.6

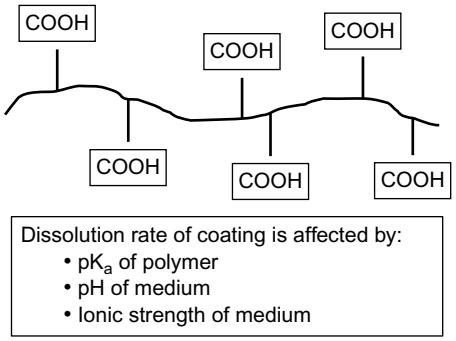


FIGURE 34.16 Structure of enteric-coating polymers.

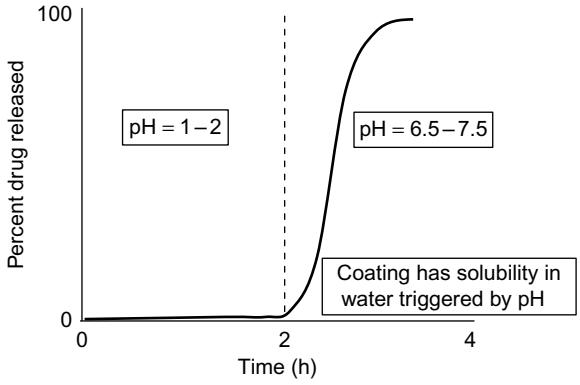


FIGURE 34.17 Typical drug release profile from products coated with enteric coatings.

34.2.4 Film coatings used for modified-release applications

As described previously, film coatings used for modified-release applications can be subdivided into two main categories, namely delayed-release (or enteric) coatings, and extended-release coatings.

34.2.4.1 Delayed-release (enteric) coatings

Enteric coatings are primarily used for the purpose of:

- Maintaining the stability of APIs that are unstable when exposed to the acidic conditions of the gastric milieu. Such API's include erythromycin, pancreatin, and the class of proton pump inhibitors, such as omeprazole.
- Minimizing the side effects (eg, nausea, and gastric irritation and bleeding) that can occur with APIs such aspirin and certain nonsteroidal inflammatory compounds.
- Creating opportunities for "night-time dosing" strategies, where the intent is to allow the dosage form to be consumed at bed-time, and permit effective blood levels of the API to be attained just prior to waking.
- Facilitating colonic drug delivery.

The functionality of enteric coatings is, for the most part, mediated by a change in pH of the environment to which the enteric-coated product is exposed. Enteric polymers remain unionized (and thus, insoluble) at low pH values, and begin to dissolve at a pH value of approximately 5.0–5.5. In addition, the functionality of enteric coatings can be greatly affected by many other factors, such as:

- The nature of the API contained in the dosage form; this is especially true when that API is ionic in nature.
- The quantity of coating applied; insufficient coating can result in ineffective gastric resistance, while too much applied coating can seriously delay drug release when the dosage form passes into the small intestine.
- The presence of imperfections in the coating (eg, cracks, "pick marks," etc.) that can also lead to reduced gastric resistance.
- The chemistry of the polymer used (especially dissolution pH and dissolution rate at a given pH).
- The influence of the in-vitro test conditions used (such as pH and ionic strength of the test medium; as well as the agitation rate used in the test).

Enteric film-coating polymers are essentially polyacids (see Fig. 34.16), and typically only dissolve in water above pH = 5.0–6.0; these polymers are selected for their ability not only to form robust coatings that adhere strongly to tablet surfaces, but also to permit rapid drug release from dosage form once it passes from the stomach into the small intestine (see Fig. 34.17).

A list of commonly used enteric-coating polymers is given in Table 34.13, and these form the basis of enteric coating formulations used in either organic-solvent-based or aqueous-coating formulations. A breakdown of coating systems specially designed for aqueous-coating applications is shown in Table 34.14.

TABLE 34.13 Examples of Common Polymers Used in Enteric Coating Formulations

Polymer	Comments
Cellulose acetate phthalate	Hydrolysis potential—high[a]
Cellulose acetate trimellitate	Hydrolysis potential—medium[a]
Polyvinyl acetate phthalate	Hydrolysis potential—low[a]
Hydroxypropylmethylcellulose phthalate	Hydrolysis potential—medium[a]
Hydroxypropylmethylcellulose acetate succinate	Hydrolysis potential—low[a]
Poly (MA-EA) 1:1	–
Poly (MA-MMA) 1:1	Relatively high dissolution pH
Poly (MA-MMA) 1:2	Relatively high dissolution pH

[a]*When exposed to conditions of elevated temperature and humidity.*
MA, Methacrylic acid; EA, Ethyl acrylate; MMA, Methyl methacrylate.

TABLE 34.14 Examples of Aqueous Enteric Coating Systems

Product	Form	Polymer
Eudragit L30D[a]	Latex dispersion	Poly (MA-EA)
Eudragit L100-55[a]	Spray-dried latex	Poly (MA-EA)
HP-F	Micronized dry powder	HPMCP
Sureteric	Formulated, dry powder system	PVAP
Acryl-Eze	Formulated, dry powder system	Poly (MA-EA)
Aquarius Control ENA	Formulated, dry powder system	Poly (MA-EA)
Aquateric	Spray-dried pseudo latex	CAP
Aquacoat ECD	Pseudo latex dispersion	CAP
Aquasolve	Micronized, dry powder	HPMCAS
CAP	Dry powder	CAP
CAT	Dry powder	CAT

[a]*Competitive acrylic products now available from BASF, Eastman, & Sanyo.*
MA, Methacrylic acid; EA, Ethyl acrylate.

While enteric-coated products have conventionally taken the form of tablets, recently a preference has been shown for coating multiparticulates because of their more consistent gastrointestinal transit characteristics. Enteric-coated capsules (especially one-piece softgels, containing garlic or fish oils used in nutraceutical applications) have also become quite commonplace.

34.2.4.2 *Extended-release coatings*

Extended-release coatings are typically completely insoluble in water, and permit the release of the API by means of diffusion through an intact coating membrane. Polymer selection, and the general formulation strategies employed in this category will very much depend on a number of factors, including a desire to create a specific type of drug-release characteristic, minimize the risk of dose-dumping, utilize processing methodologies that already exist within the company, and prepare a unique dosage form that enables the manufacturer to take a proprietary position with respect to dosage-form presentation.

In order to meet these requirements, critical objectives include the need to:

- Achieve the target drug-release characteristics in a reproducible manner;
- Ensure that drug-release characteristics are insensitive to expected variations in raw materials and coating-process conditions;
- Confirm that the coating formulations (and associated coating processes) are essentially uncomplicated, and facilitate scale-up from the laboratory into production; and
- Ensure that the final product is stable and does not exhibit time-dependent changes in drug-release characteristics.

When it comes to designing extended-release products, these may take a number of forms, namely tablets, prepared as film-coated tablets, film-coated mini-tablets, or compacted, film-coated particulates, as well as capsules, prepared as encapsulated, film-coated particulates, or film-coated capsules.

Coated multiparticulates often form the basis for extended-release products when an applied coating is the main mediator of drug release because of the desire to minimize dose dumping, and facilitate more consistent gastro-intestinal transit times. Such coated multiparticulates (see Fig. 34.18) can take the form of drug-loaded pellets (nonpareils), granules (irregularly shaped granules or granules that have been spheronized), drug crystals, and finally drug/ion-exchange-resin complexes.

In terms of designing extended-release coating formulations, the types of polymers used are often selected for their ability to exhibit similar robustness characteristics as those described for enteric coatings, but also for their permeability characteristics that allow the drug to be released at a controlled rate that is consistent with the biological requirements for that API (see Fig. 34.19).

A list of common polymers used for extended-release coating applications is shown in Table 34.15, while examples of aqueous-coating systems are shown in Table 34.16.

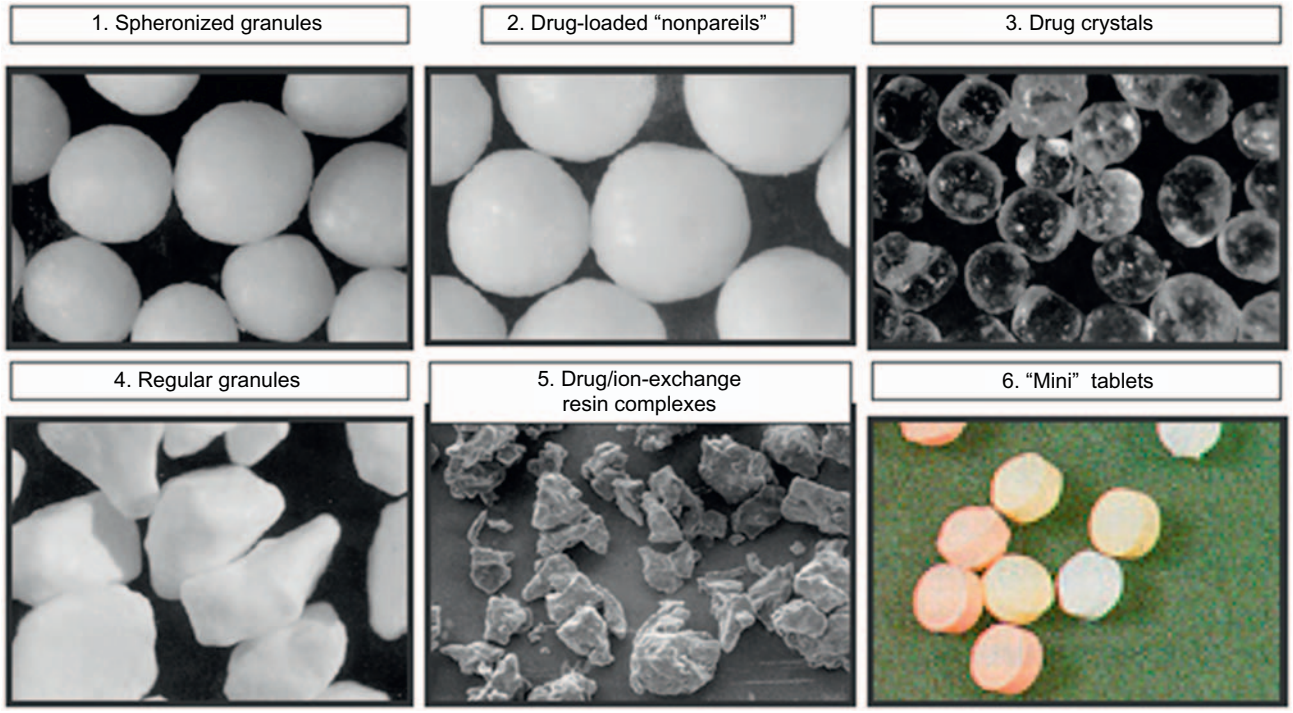


FIGURE 34.18 Examples of multiparticulates.

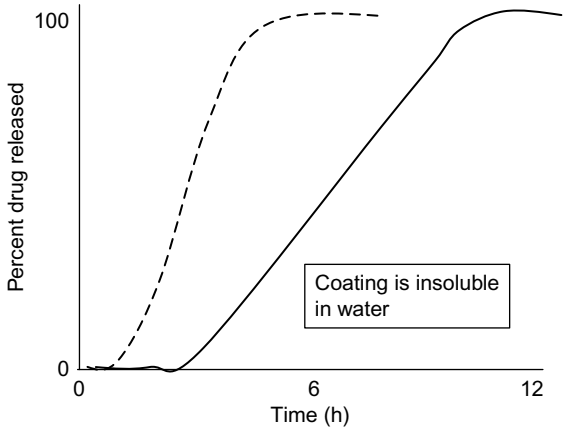


FIGURE 34.19 Typical drug release profile from products coated with extended-release coatings.

TABLE 34.15 Examples of Common Materials Used in Extended-Release Coating Formulations

Coating material	Membrane characteristics
Fats and waxes (beeswax; carnauba wax; cetyl alcohol; cetostearyl alcohol)	• Permeable and erodible • Permeable and soluble (at high pH)
Shellac	• Permeable and soluble (at high pH)
Zein	• Permeable and water-insoluble
Ethylcellulose	• Semi-permeable and water-insoluble
Cellulose esters (eg, acetate)	• Permeable and water-insoluble
Acrylic ester copolymers	

TABLE 34.16 Examples of Aqueous Extended-Release Coating Systems

Product	Polymer used	Comments
Surelease	Ethylcellulose	Plasticized aqueous polymer dispersion, addition of lake colorants should be avoided
Aquacoat	Ethylcellulose	Pseudo latex dispersion, plasticizer must be added to facilitate film formation
Eudragit NE 30D[a]	Acrylic copolymer	Latex dispersion; no plasticizer needed unless it is necessary to improve film flexibility
Eudragit RL 30D[a]	Acrylic copolymer	Aqueous polymer dispersion; no plasticizer needed unless it is necessary to improve film flexibility
Eudragit RS 30D[a]	Acrylic copolymer	Aqueous polymer dispersion; no plasticizer needed unless it is necessary to improve film flexibility

[a]*Also available as Kollicoat Systems.*

The performance of extended-release coating formulations is often defined by:

- The nature of polymer used, especially in terms of its chemistry (which can greatly influence coating permeability) and molecular weight (which has a significant effect on coating robustness).

- Presence of additives, such as:
  - Plasticizers (which can greatly influence coating robustness, as well as the extent of film formation when using aqueous polymer dispersions);
  - Anti-adhesive agents (which are often used to reduce agglomeration in the case of multiparticulates, as well as reduce imperfections in the coating caused by picking and sticking);
  - Colorants which, although used for identification purposes, can influence film permeability; and
  - Water-soluble pore-forming agents, which are typically added to the formulation as a deliberate strategy to modify the permeability characteristics of the applied coating.
- Coating thickness which is, in turn, influenced by the:
  - Surface area to be covered,
  - Quantity of coating applied,
  - Uniformity of distribution of coating, and
  - Coating process efficiency.

While extended-release coating formulations have typically been designed as polymer solutions in organic solvents, there has been significant interest over the past 20 years in using aqueous-coating systems. Irrespective of the formulation approach used, it is critical that the formulation of the coating enables the appropriate drug-release characteristics to be achieved in a consistent manner, the membrane obtained should be structurally sound, the coating systems are readily adaptable to existing coating-process technology, and the coating formulation and processes are sufficiently optimized to prevent the occurrence of time-dependent changes in drug-release characteristics.

Meeting these objectives can be more challenging when using aqueous polymer dispersions because it is necessary to ensure that coalescence of the coating during film formation proceeds to an endpoint where membrane porosity is eliminated, this completeness of coalescence is either achieved during the coating process, or by means of a short curing step performed at the conclusion of that process, and further gradual coalescence, over an extended period of time, is avoided.

Thus the coating formulation used must be suitably optimized in terms of its ability to achieve completeness of film coalescence in the particular coating process to be employed. This usually means that the MFFT (which is closely related to the glass-transition temperature of the polymer employed, either in its natural state, or as modified by the presence of a plasticizer) is matched to the typical temperature conditions used in the process. In addition, the coating process conditions employed must also be optimized to ensure that they enable the correct temperature conditions to be achieved (to facilitate film coalescence).

34.3 DESIGN AND DEVELOPMENT OF FILM-COATING PROCESSES

34.3.1 General introduction to coating processes and equipment

34.3.1.1 Overview

A film coating is usually applied onto tablets until they reach a targeted weight range, often expressed as a percentage weight gain. A typical weight gain for an esthetic coating is 2–4%, whereas, for clear coatings, the weight gain may be as little as 0.5–1.0%. For controlled-release applications, the quantity of the coating required is directly related to its thickness. A thickness of 30–50 μm (this translates to 4–6 mg polymer/cm^2 of tablet surface) is usually sufficient to provide a satisfactory enteric coating. It will require a greater weight gain to achieve a controlled release film of a suitable thickness because a batch of smaller tablets contains a greater total tablet surface area.

The reasons for selecting a specific film-coating process are essentially the same, regardless of whether the process is performed in a perforated coating pan or a fluid-bed coating process. In general, the major criterion for deciding between either a coating pan or a fluid-bed process is often related to the size and shape of the product to be coated. A rule of thumb is that if the product is approximately 6 mm or less in diameter, then the preferred equipment for applying the coating is the fluid-bed coater, mainly because when attempting to coat small particles in a perforated coating pan (where process air is drawn down through the tablet bed under a negative pressure) the pressure drop across the product bed can drastically reduce the process air volume, a factor which will reduce the evaporative capacity of the pan. This, coupled with the close contact of the product in the spray zone (increasing the risk of agglomeration), makes the coating of small particles in the coating pan less desirable. When attempting to coat larger size products (such as tablets) in a fluid-bed process, often the capacity of the fan is insufficient to fluidize such a product, and even when successful fluidization is achieved, product attrition may be excessive.

34.3.2 Batch coating systems

Batch coating of solid oral dosage forms requires a process where the equipment involved not only acts as a container, but also provides constant motion and facilitates the evaporation of the coating solvent. In a pan-coating process, the container will typically be cylindrical and will rotate on a horizontal axis. The coating formulation will be sprayed onto the surface of the constantly moving product in a finely atomized

state, typically achieved using a spray gun. Droplets from the spray gun are composed of a solution (or suspension), which is a mixture of a liquid carrier and solids. The principle of this type of coating is to apply the solution in such a manner that a droplet from the spray gun will land on the product and spread. The temperature of the product will then rapidly evaporate the liquid from the suspension, leaving only the dried solids. The coating solution/suspension viscosity must be such that any evaporation that occurs between emergence of droplets from the spray nozzles until the time they impact the surface of the product being coated, these droplets will retain sufficient fluidity to enable them to spread into a uniform coating before solvent evaporation is essentially completed. This will leave a dry coating, consisting of the solids that were in the solution, on the surface of the product.

In order to produce a uniform coating, each tablet must have an equal probability of passing through the spray zone throughout the course of the coating process. In order to facilitate such uniform distribution of the coating, coating pans typically have some type of mixing baffle that not only facilitates the tumbling action of the product being coated, but ensures that product is moved back and across the coating pan. Finally, in order to promote efficient drying within the process, the coating pan must not only facilitate the application of process air across the surface of the tablet bed, but also ensure that this process air can pass through the tablet bed to complete the drying process.

Coating pans described in this section are typically considered to be closed systems, where the environment outside of the coating pan is completely separated from that within the coating pan, so that material is contained within the process and drying efficiency is maximized.

Overall, the challenge in film coating is to ensure that the coating is uniformly applied across the surface of the product to be coated, and that process thermodynamics (which defines the balance between the rate at which the coating liquid is applied and the rate at which the coating solvent is removed) are controlled in such a manner that prevents both tablets from being overwetted, such that they stick together, and that overall process efficiency (the relationship between the amount of coating material delivered into the process and that which actually ends up on the tablets) is maximized.

34.3.3 Continuous-coating systems

A continuous-coating system differs from a batch system in that tablets are continually fed into the process, pass through the spray zone in a short period of time (typically 15–20 minutes), and then are continually emptied from the process in the former. In a batch process, the tablets reside inside the coating pan for the duration of the coating process (potentially 1–4 hours depending on batch size), and only emerge once the entire process is completed. Despite these differences, the coating principles, as well as process thermodynamics, are essentially similar. A typical batch coating process is shown in Fig. 34.20, while a continuous process is exemplified in Fig. 34.21.

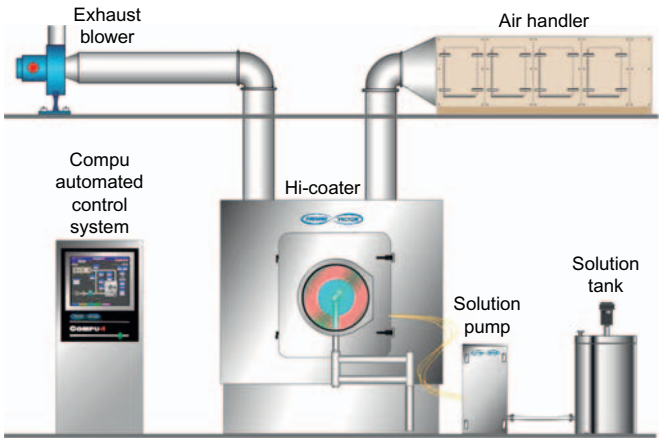


FIGURE 34.20 Example of a coating system for batch processing.

34.3.4 System components

The peripheral equipment for a coating system for either the batch system or the continuous-coating system will be very similar. Figs. 34.20 and 34.21 show equipment components typically used with both batch- and continuous-coating systems.

34.3.4.1 Overview

Process air is drawn in through an air handler where it will be heated, dehumidified (or humidified, depending upon requirements), and filtered before going to the coating pan. The process air enters the coating pan and is exhausted via a plenum located below the tablet bed, to a dust collector and a blower. Finally, the exhaust air is discharged into the environment (when organic solvent-based coating systems are used, the exhaust air will pass through a solvent recovery unit or an incinerator before being discharged). Some organic solvent systems may use a closed-loop system where solvent is condensed from the air stream before being fed back into the inlet process air stream. A closed-loop system has the advantage of using process air, which is conditioned to suit the operating parameters required for the process, thus requiring less energy input for the operation.

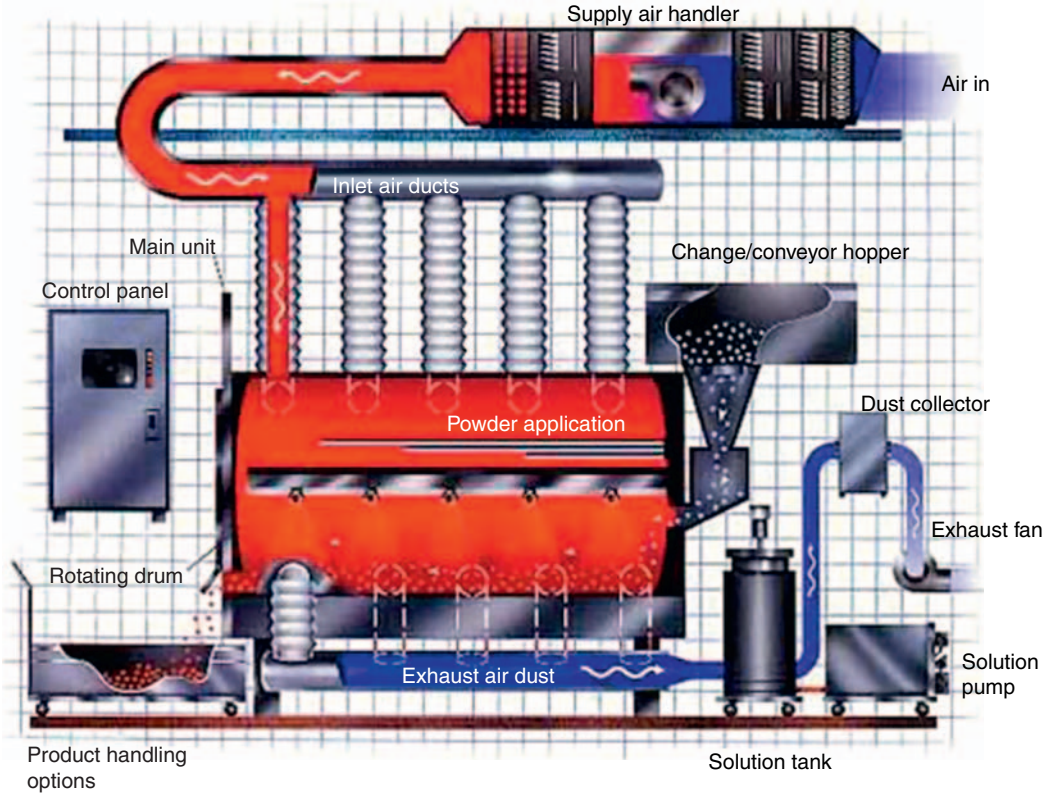


FIGURE 34.21 Example of a coating system for continuous process.

34.3.4.2 *Pan units*

34.3.4.2.1 Comparison of batch-type coating pans

The coating pan can be either fully or partially perforated around the cylindrical surface of the drum, which allows process air to be drawn through the tablet bed, ensuring more effective drying. An example of a partially perforated coating drum is shown in Figs. 34.22 and 34.23, and displays an integrated exhaust plenum on the outside of the drum. A fully perforated coating pan is shown in Figs. 34.24 and 34.25.

As the coating drum rotates, the product being coated (eg, tablets) tumbles in the drum, always remaining near the bottom of the pan (as indicated in Figs. 34.23 and 34.25). Partially perforated coating pans are constructed with exhaust plenums welded to the outside of the pan. Process air enters the coating pan through a mouth ring at the front opening of the pan, or through the back of the pan, and is drawn out through the exhaust plenums built onto the exterior of the cylindrical portion of the drum, see Fig. 34.23. The perforations in the coating pan are located over these plenums. This arrangement forces all of the process air to pass through the tablet bed. There is greater

FIGURE 34.22 Exhaust plenum on the exterior of a partially perforated coating drum.

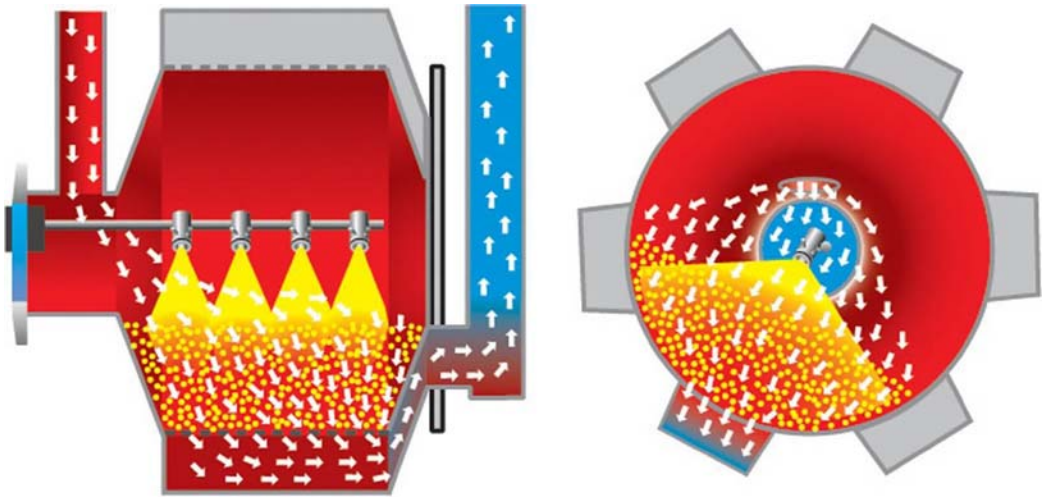


FIGURE 34.23 Air-flow diagram and chamber design of a partially perforated pan.

FIGURE 34.24 Exterior of the fully perforated coating drum.

pressure drop in the path of the process air (compared with the fully perforated pan) because of the plenum system and the tablets. The path of the process air through the fully perforated pan is shown in Fig. 34.25.

Coating can be performed with any of these pans if the auxiliary equipment (blowers, heater, filters, etc.) is configured and specified properly for the physical layout. Some coating-pan equipment includes auto discharge. The mixing baffles can be configured in these pans such that when the pan is rotated in a reverse direction, the product will be discharged out of the door. Some coating pans have a trap door on the product drum so that the product can be discharged into product trays placed under the units. Several of the manufacturers will also supply product-loading totes, fitted with chutes, that will allow the product to be loaded from above the coating pan instead of through the drum opening.

Coating pan capacities for laboratory size units range from 0.5 to 110 L. Capacities for production-size coating pans range from 90 to 1300 L. The laboratory size units are used to perform feasibility studies and for the purposes of product and process development. An example of laboratory-size units is shown in Fig. 34.26. In general, each manufacturer makes sizes of coating pans that are comparable, although the trend now is to provide lab-scale coating equipment with multiple pan-size inserts.

Each manufacturer will provide specifications for general operation of their coating pans. Typically, they will recommend and/or furnish the peripheral equipment required to complete a system.

34.3.4.2.2 Comparison of continuous-coating equipment

An example of a continuous-coating unit is shown in Fig. 34.27. The product is moved through a drum (that represents a stretched version of a batch-coating pan) at a rate that allows the tablets to be completely coated as they pass through the coating pan. Mixing baffles may be present to help move product from one end of the coating unit to the other end. The application of coating material within a continuous coating pan is essentially the same as that in a batch side-vented coating pan. Continuous coating pans are available from Thomas Engineering, O'Hara Technologies, Driam, Bohle, and GEA.

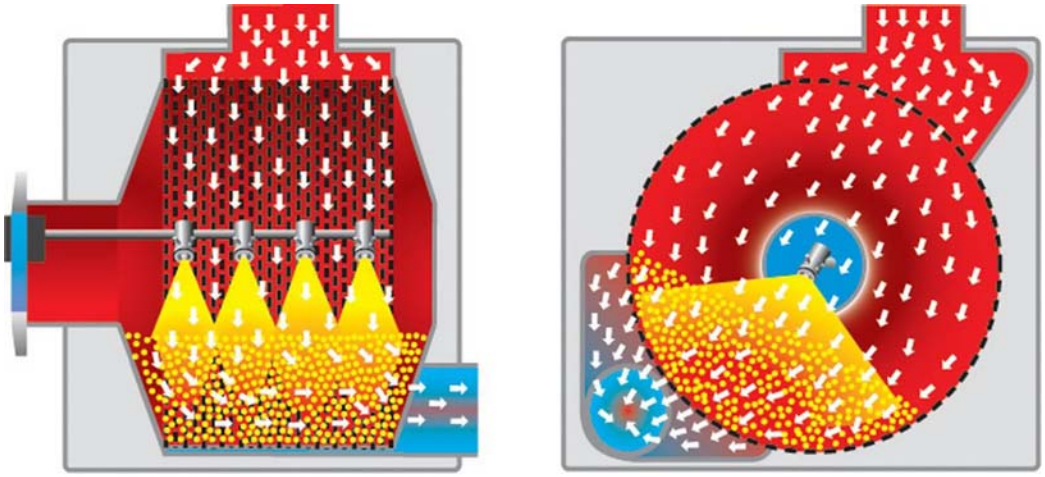


FIGURE 34.25 Air-flow diagram and chamber design of a fully perforated pan.

FIGURE 34.26 An example of a laboratory-size coating pan.

FIGURE 34.27 Example of a continuous-coating pan.

34.3.4.3 *Process air equipment*

All coating-pan equipment, including continuous-coating process equipment, requires conditioned-process air to enable solvents (aqueous or organic) to evaporate quickly. The process air generally should be dehumidified, cleaned, and heated. An air handler is used to perform these functions. A typical air handler is made up of different functional sections as shown in Fig. 34.28.

The first section will generally be composed of two filters: a bird screen followed by a prefilter. The prefilter is typically rated for 30% efficiency. The next section of the air handler is a preheat unit, followed by a dehumidification unit. The preheat section is placed before the dehumidification unit to protect any chilled water coil (or refrigeration coil) from freezing. Three methods of dehumidification are used: chilled water, refrigeration, or a desiccant dryer. The desiccant dryer is generally used to obtain the lower dew points. If humidification is required (ie, during winter when the air is very dry), this is accomplished by injecting steam into the process air stream. A heating coil is required to heat the process air.

The air handler shown in Fig. 34.28 is a face and bypass unit, which allows heated and room-temperature

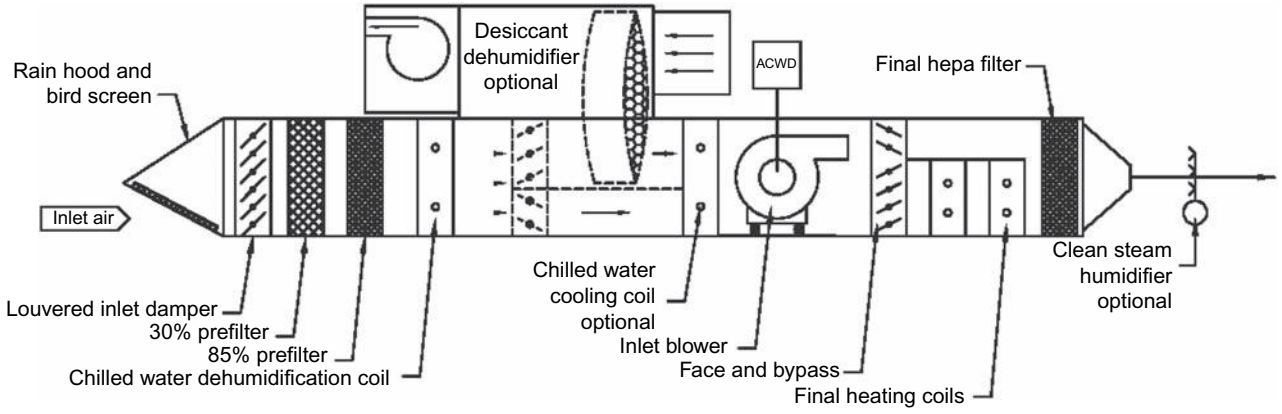


FIGURE 34.28 Example of an air-handling system.

air streams to be blended to achieve the required set-point temperature, as well as to allow temperature changes to be made quickly and facilitate better control of temperature. A conventional (nonbypass type) air handler will require some time to cool down because the heater coils remain in the process air stream after the heat has been stopped. In addition, as heat (either electric or steam) is provided, there tends to be a greater risk of oscillations in temperature occurring.

The finished product is normally cooled before being removed from the coating pan after spray application of the coating is completed. The face and bypass unit can achieve a shorter production time because a heating coil does not have to cool down before cooling the product. The air handler comprises a fan (blower), which must be rated for the proper quantity of process air to facilitate heat transfer to the product. The final stage of an air handler should be a high-efficiency particulate air (HEPA) filter to provide clean air with 99.97% efficiency.

The schematic shown in Fig. 34.28 illustrates the typical components of an air-handling system. The duct work between the HEPA filter and the coating pan should be stainless steel to preserve the integrity of the clean, process air. The quantity of process air is generally recommended by the equipment manufacturer for the size of the equipment. It is desirable to use the maximum amount of process air possible to achieve the maximum heat transfer and drying efficiency. However, caution must be exercised when increasing the air flow within the system because pressure drop increases exponentially and turbulence can be generated that will affect the spray patterns inside the coating pan. The inlet blower should be capable of providing sufficient air pressure at the desired quantity of air to bring the pressure in the coating pan to atmospheric pressure or slightly negative (up to −100 mm of water column), when the exhaust blower is operating and product is loaded in the pan. The exhaust blower must provide a vacuum pressure at the required air flow that will overcome the losses created by the different components between the coating pan and the exhaust blower.

The outlet process air duct should be connected to a dust collector to collect solids that did not adhere to the product or dust that is generated when product is loaded into the coating pan. The duct work before and after the dust collector, and the dust collector itself, must be able to withstand the amount of vacuum that will be created by the exhaust blower. The dust collector also must be designed with an explosion vent that will relieve the pressure in case of a dust explosion. When film coating is performed using an organic solvent, the entire system must be designed either to contain an explosion, or allow the explosion to be safely vented. Typically, the equipment manufacturer will recommend a system that will satisfy the requirements for a specific type of solvent.

As mentioned previously, the partially perforated coating pan generally creates greater pressure drop, thus requiring a high-pressure blower; whereas a fully perforated coating pan produces a lower pressure drop, but requires a larger quantity of process air. Finally, the exhaust blower should be capable of generating the recommended air flow for heat transfer and producing a vacuum of 25−100 mm of water column inside the coating pan. In areas where the noise of the blower cannot be tolerated, especially on the systems that require a high pressure blower, a silencer should be used after the exhaust blower. Most manufacturers of coating pans will provide the above peripheral equipment as part of a turnkey system. Some systems will have a bypass section that allows the process air to bypass the coating pan, typically being used where the coating pan door must be opened for several minutes, thus allowing cool air to enter the coating pan. If the temperature probe, which is used for the control loop,

is in the exhaust duct work, the control loop will call for a higher temperature. When the coating pan door is closed, the inlet process air temperature will be considerably higher, which in turn can cause the product temperature to exceed the set limits. The bypass allows the process air temperature to remain constant. The bypass also must have a pressure drop to prevent the air flow from increasing, thus causing a temperature change during bypass conditions. Utilizing the bypass will then allow the temperature control loop to hold the process air at a constant temperature during the time the airflow is not passing through the coating pan.

34.3.4.4 Spray systems

The design and operation of the spray-application system is a critical element of the coating process. In general, two types of spray systems are used, an open-loop system or a recirculation-loop arrangement. An open-loop arrangement is typically used with both hydraulic spray guns (atomization is generated through a high liquid-line pressure) and pneumatic spray guns (atomization is generated by a compressed air stream that impinges on the liquid stream as it emerges from the nozzle), while a recirculation loop is typically used only with pneumatic spray guns. Generally, a recirculation loop provides more flexibility, but requires more equipment and tubing to configure. When the spray is off, the coating formulation is recirculated back to the solution tank, providing mixing in the tubing to prevent precipitation of solids out of solution. A separate agitator in the solution tank should be employed for coating formulations prone to precipitation. To use the recirculation loop, a valve in the spray gun and a valve in the recirculation loop must be provided to change from spray to recirculation. A flow-measuring device is not needed for a positive displacement pump, but is desirable. A separate pump can be provided for each spray gun or a single pump connected through a manifold can feed all spray guns. The manifold can create different pressures at each gun due to the pressure drops in the manifold tubing, which in turn will cause different droplet sizes and distributions at each gun; a manifold system also has the additional disadvantage of causing extra liquid to be fed to the remaining spray guns should one or more nozzles become blocked. Many spray systems use a multiple pump configuration to prevent the pressure drop problem.

34.3.4.4.1 Comparison of different spray guns

Two types of spray guns are typically used for coating: either pneumatic or hydraulic. A third type is an ultrasonic, which has had limited success for coating processes. For most circumstances, the use of pneumatic spray guns dominates the industry today because of the preference for using aqueous-coating systems.

34.3.4.4.2 Pneumatic spray guns

The pneumatic gun sprays liquid out through an orifice, which is then subjected to a stream of compressed air that breaks the spray into droplets. The size of the droplets, the distribution of droplets, and the type of pattern is determined by the quantity of compressed air flow and the physical configuration of the spray gun. An illustration of a pneumatic spray gun is provided in Fig. 34.29, and shows the different ports for the compressed air supply and solution feed.

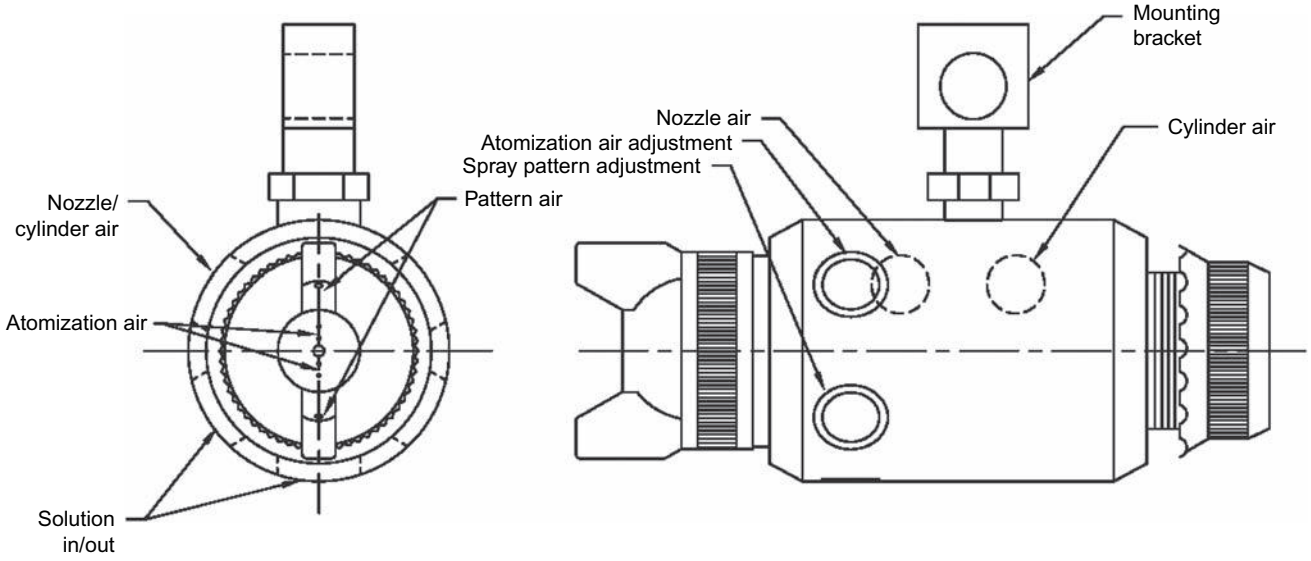


FIGURE 34.29 Example of a pneumatic spray gun with separate adjustments for atomization air and pattern air.

TABLE 34.17 Common Spray Gun Manufacturers

Manufacturer	Address	Trade name	Types of spray guns
Binks Sames Corporation	Geneseo, IL 61254	Binks Guns	Pneumatic
Freund (Vector Corp.)	Marion, IA 52302	Vector or Freund Guns	Pneumatic
Schlick	Coburg, GERMANY 96540	Schlick Guns	Pneumatic Hydraulic
Spraying Systems Co.	Wheaton, IL 60189	Spraying Systems	Pneumatic Hydraulic

Table 34.17 provides manufacturers of spray guns.

Solution and air orifice sizes on the spray guns may need to be changed for different types of coating suspensions but, in general, a single orifice size will suffice for most commonly used coating suspensions.

34.3.4.4.3 Hydraulic spray guns

A hydraulic spray gun uses pressure on the liquid feed lines to force the coating liquid through an orifice where, under the influence of the pressure drop created at the nozzle, the liquid is sheared into droplets. The orifice must be changed for different types of spray patterns and spray distributions. Hydraulic spray guns have limited use in coating pans because the spray distribution is considerably wider than for pneumatic spray guns. The typical pressures required to maintain effective atomization also often result in spray rates being too high for aqueous applications; however, this type of nozzle can be very beneficial when using organic solvent-based coating formulations.

In all types of spray guns, as the spray rate is changed, the pattern, droplet size, droplet distribution, and droplet velocity will also change. The pneumatic gun provides more flexibility because the atomizing air can be varied to obtain the desired spray droplet characteristics without changing the nozzle. Ultrasonic spray guns produce relatively small droplet size over the wide range of spray rates, and have the advantage in that the larger orifices they employ result in less plugging of the nozzle. Both the hydraulic and pneumatic spray guns are subject to plugging, causing disruption of the droplet size distribution and the spray pattern. The hydraulic gun is more prone to plugging than the pneumatic gun because the orifice openings are smaller.

34.3.4.4.4 Solution delivery pump

The open-loop arrangement described previously for the spray-application system is generally used with pumps, such as a gear pumps, a lobe pump, or peristaltic pumps. These pumps are capable of supplying pressures of 1–7 bar gage pressure to the spray nozzles. Peristaltic pumps are most commonly used because of the ease of cleaning and lower cost. Centrifugal pumps are normally not used because of pressure requirements and the need for a flow measurement service due to the characteristics of the pump.

The selection of the pump is determined by the type of spray guns used. The hydraulic spray guns will require 2–4 bar of solution pressure at the desired flow rates. The pneumatic spray guns will require 0.05–1.0 bar at the desired flow rate. The ultrasonic spray guns require only pressure to supply the solution to the spray head (typically less than 0.7 bar). Normally, pumps that operate at lower pressures create fewer problems relative to pump wear and sealing of connections to the pump and spray guns. To ensure that proper flow and pressure is provided to each spray gun, many of the spray systems incorporate a separate pump (or pump head driven from a common drive system) for each spray gun. A single pump or a single tubing distribution system, through a manifold to the spray guns will result in different pressures and flow rates at the spray guns. When selecting a pump, the location from which the coating solution is being drawn must be considered. Some coating formulations may require additional pump pressure due to a high vacuum pressure on the inlet side of the pump. When using more than one pump, the inlet tubing must be sized for the total number of pumps and the vacuum loss due to the size of the inlet tubing.

The connections for the spray system are very critical and must be installed with care. If there is a leak in the tubing on the inlet side of the pump, air bubbles are created in the suspension which travel to the spray heads and cause distortion of the spray pattern and droplet size distribution. Any leaks in the tubing, piping, or connections will result in dripping onto the product because all spray guns are mounted inside the coating pan over the product; causing twinning (tablets stuck together) or picking (tablets stuck together and separated).

34.3.4.4.5 Delivery control

The application of the coating liquid onto the tablets can only be consistent if the rate at which the liquid is being delivered to each spray nozzle is also consistent. For this reason, effective control of the flow of coating suspension must be exercised. In the past, assumptions have been made that if a positive displacement pump is used, then the rotational speed of the pump shaft (rpm) can be used to determine the flow rate. However, variations in the spray gun outputs can

occur because of pressure changes in the system and systems where more than one pump is used. Therefore it is advisable that some form of measurement and control be used for the suspension flow. The most common form of measurement is a mass-flow meter. Spray systems that use a manifold for the spray guns can be configured to use a single pump, but a mass-flow meter should be used at the outlet of the pump. Spray systems that use multiple pumps feeding individual guns should incorporate a mass flow meter in the common-source, line-feeding the pumps. There are several types of mass-flow meters that have been used including vortex-flow meters, Coriolis mass-flow meters, and turbines. Turbines are not typically used because they cannot be cleaned properly without removing them from the coating suspension line. Differential pressure is not used because this requires an orifice plate which generates more pressure drop for the spray system, thus increasing the work load for the pump. These flow devices will normally have to measure mass flow over a range of 100−1000 mL/minute. An alternative means to measure the application of the coating is to use a weight-loss system, which involves putting the solution tank on load cells (weight measuring sensors) and then, knowing the density of the suspension, the amount of material applied can be determined as a reduction in the tank weight. This weighing system must be capable of having the tare weight set to zero before starting application of the coating liquid.

34.3.4.4.6 System controls

System controls may take one of several forms and should be chosen based on the tolerance required for the product, the operator's technical knowledge level, and the extent of control desired. There are basically three types of systems:

- Manual: All functions are selected and controlled by an operator;
- Semiautomatic: These systems use a combination of manual and automatic systems; and
- Automatic: This system requires minimal operator input during the coating process.

Typically, the design of a coating process is based on the equipment used, and the process conditions that are deemed critical to the overall success of that process. Ideally, for any given process, these conditions should be held constant:

- Process air flow
- Inlet process air temperature
- Average static pressure in the coating pan
- Humidity (or dew-point) in the process air going into the coating pan
- Coating pan speed
- Spray rate of the coating solution

All three control systems mentioned previously will require certain parameters to be measured in order to make any coating system operate with consistent results. The manual system will have only instruments that provide local readouts, which are used by the operator to adjust the controls. Both types of automation will require outputs that can be input into a programmable logic controller (PLC) or computer.

These measurements represent the minimum required for any of the coating systems to provide a consistent coating operation:

- Inlet process air temperature
- Exhaust process air temperature from the coating pan
- Process air flow volume or flow rate in the inlet duct to the coating pan
- Pump speed or suspension spray rate.
- Coating pan rotation (RPM)

These controls are required for each of the systems:

- Power to the system
- Temperature controller
- Air-flow control (manual or motorized dampers, or blower motor speed control)
- Solution volume or rate control (mechanical or pump motor speed control)
- Coating pan rotational speed (mechanical or pan motor speed control)

Each of these items for either measurement or control will be analog in nature, either by an analog meter or an analog signal to a PLC or computer.

Other measurements that may be included with a coating pan system are:

- Process air temperature after the preheater
- Dew-point temperature after the dehumidification or humidification
- Differential pressure across the inlet filter to the air handler
- Differential pressure across the HEPA filters in the inlet air duct
- Differential pressure across the coating pan
- Static pressure in the coating pan
- Differential pressure across the dust collector
- Static pressure after the air handler
- Static pressure at the inlet of the exhaust blower
- Product temperature
- Main compressed-air pressure
- Solution tank weighing system
- Solution pressure
- Solution mass-flow meter
- Compressed air flow rate to pneumatic spray guns

34.3.5 General characteristics of the pharmaceutical coating process

34.3.5.1 Typical process steps

If this is the first time that a particular product has been coated, it is generally necessary to estimate the approximate batch size to use. Coating pans are usually rated for a specific brim volume capacity. Brim volume can be defined as volume capacity, if loaded to the very bottom of the pan opening. The initial pan load is usually somewhat less than brim volume in order to prevent product from spilling out of the pan opening should the product volume increase during the coating process or if the product movement should change. As a starting point, the brim volume capacity should be multiplied by 95% in order to estimate the starting fill volume; this estimated volume can then be multiplied by the product's bulk density (in g/mL) to determine the approximate maximum batch capacity. Product bulk density can be approximated by determining the weight of a 1-L volume of tablets. Batches of 70–75% of the maximum can usually be coated without a problem. For smaller batch sizes, it is necessary to examine product movement to verify that it is adequate, and also determine that the exhaust air plenums (in a fully perforated pan) are completely covered. Reduced height baffles can be used to improve product movement with smaller pan loading. In cases where the exhaust plenum is not completely covered, it may also be necessary to use a restrictor plate to reduce the exhaust opening, thus preventing air from being drawn around the tablet bed rather than through it.

For continuous coating pan systems, the throughput is controlled by a combination of the pan drum angle, pan speed (rpm), and tablet input rate. A higher throughput is achieved by increasing one or more of these factors. Either one of these changes causes the product to flow more rapidly through the pan, thus increasing the potential production rate.

34.3.5.1.1 Coating pan set-up

The typical coating process consists of several different steps. The first step is to verify proper operation of the spray delivery system, which must be calibrated to deliver a consistent gun-to-gun delivery of the coating solution, especially when manifold systems are used. Calibration can be accomplished by individually collecting and weighing the solution from each spray gun for a set interval of time. The calibration procedure should determine if the overall delivery of solution is accurate (if it matches the theoretical rate) and also confirm that the variation in solution flow between guns is minimized. If there is an unacceptable variation (usually $> \pm 5.0\%$), then the system will need to be adjusted so that the pressure drop between guns is the same. Most spray guns are equipped with a needle valve-adjustment cap that controls the clearance between the spray needle and the liquid nozzle. Adjustment of the position of the needle valve will in turn facilitate adjustment of the fluid delivery rate through a particular spray gun. After making an adjustment to the spray guns, the solution collection should be repeated. When uniform solution delivery has been achieved, it is necessary to calibrate the nozzle air volume to the spray guns. Some spray guns have a common line for providing atomization air and pattern air. For these systems, the air volume is achieved by setting the desired nozzle air pressure. Other guns have separate controls for the atomization and pattern air, and a mist checker or flow meter for display of the actual air volume. For these systems, a needle valve (located on the spray gun body) is usually adjusted to set the atomization air volume. Once the atomization air volume has been set for all of the guns, a second needle valve is adjusted to the desired volume for pattern air. After these air volumes have been set, it is possible to verify the approximate dimensions of the spray pattern by passing a hand through the air stream. This spray pattern should be checked at a distance that is equal to the distance between the tablet bed and the spray guns.

34.3.5.1.2 Loading/charging

After proper set-up of the pan and spray equipment has been confirmed, the pan can be loaded with product either through the front of the pan (pan mouth ring) or through a discharge door located on the flat of the pan (if it is so equipped). During the loading of the pan, the exhaust air is usually turned on to eliminate or minimize exposure to irritating and potentially hazardous dusts. It is a good idea to minimize the distance that the product is allowed to drop during charging because this will minimize or eliminate tablet breakage. It is usually recommended that the pan be jogged occasionally during loading to move product towards the back of the pan.

34.3.5.1.3 Preheat/de-dusting

After loading, the next step is to preheat the product to the desired process temperature. While the product is being preheated, it is also de-dusted by the process airflow. This is a good time to circulate coating solution through the spray guns (if a recirculation system is used). Verification that the spray guns are positioned correctly can be accomplished at this time, checking both the angle to and distance from the tablet bed. During the preheat step, the product can either be rotated continuously at a very slow speed (preferred)

or jogged intermittently. This will minimize attrition while ensuring that the product is uniformly heated. For heat-sensitive products, it will be necessary to jog on a more frequent basis to prevent overheating the upper surface of the tablet bed. Tablets are generally preheated to a specific product or exhaust temperature. The product and exhaust temperatures are typically not the same, though they tend be fairly close during the coating process. Once the exhaust air or product has been heated to the desired starting temperature, application of the coating liquid can be initiated.

34.3.5.1.4 Application of a seal/barrier coat

In a few rare cases, it may be necessary to apply a barrier-coating after the application of the primary film coating. The need for a barrier/seal coat typically occurs when there is an interaction between the film polymer and the product substrate (eg, interaction between an enteric polymer and tablets containing an alkaline drug, or interaction between an acid labile drug and an enteric polymer, which is acidic). In such cases, a seal coat of an inert film polymer, such as HPMC prior to the enteric polymer coating, may be used. A second example occurs when there is an interaction between the coating solvent and the product (ie, an aqueous coating and an effervescent product). To remedy this type of problem, a seal coating with an alternative solvent may be used. It is possible to coat, for example, some effervescent products with aqueous solutions if the spray rate for the initial phase of the coating is reduced, process temperatures are substantially increased, and the solids content of the coating suspension is high (approximately 20% w/w).

34.3.5.1.5 Application of the film coating

Prior to commencing the spray cycle, the pan should be set to the desired rotational speed. If a frequency drive is used, it is important to allow approximately 5 seconds for the pan to achieve the rpm set-point. Once the product has reached the required speed, the spray can be started. The initiation of the spray will cause the product and exhaust temperatures to drop slightly due to evaporative cooling. If these temperatures drop below the desired range, the inlet temperature will need to be increased. If the coating system has an inlet air handler with some type of humidity control, there will be a very consistent batch to batch correlation between the inlet temperature necessary to achieve the desired exhaust and/or product temperatures. Some systems automatically control the inlet temperature to maintain the required exhaust temperature. The time required for the coating cycle is determined by a number of factors that include: desired coating weight gain, coating process efficiency, coating solution solids level, spray rate, and the size of the spray zone. The end of the spray cycle is usually controlled by either time or the application of a set quantity of coating solution because these factors tend to be constant for a particular product. If the spray cycle is controlled by solution quantity, this can be accomplished through the use of either a mass-flow meter or a mass balance. Either of these methods will provide a totalization of the solution applied.

34.3.5.1.6 Gloss coat

After the primary coating has been completed, a dilute over-coating may be applied to prevent the tablets from blocking or sticking (such as with some of the aqueous dispersions that are thermoplastic) or to provide a higher film gloss. If the purpose is to provide a higher gloss, a dilute HPMC solution may be used. The process temperatures may be reduced slightly to reduce the amount of spray drying that occurs.

For some of the aqueous coating dispersions, it is recommended that the product be maintained at a slightly elevated temperature to fully cure the coating and provide a stable release profile. After coating, the product should be cooled prior to the application of a powdered wax (if one is to be applied to improve tablet flow and gloss). Cooling the product after coating also minimizes potential problems due to heat instability of the active (since product temperature will rise after the spray is turned off due to loss of the evaporative cooling effect. A product is typically cooled to an exhaust or product temperature of 25–30°C before discharging from the pan.

34.3.5.1.7 Wax addition

After cool-down of the product, a powdered wax may be used to provide a higher tablet gloss. The waxes typically used are either a carnauba wax, a combination of carnauba and beeswax, or rice wax. A small quantity of wax (5–10 g per 100 kg of tablets) is applied to the rotating tablet bed. The tablets are allowed to rotate for approximately 5 minutes with no process air. After 5 minutes, the tablets are rotated for an additional 5 minutes with the process air on. This allows any excess wax to be exhausted. A canvas-lined pan is not necessary for the wax application, although it may provide a slightly higher gloss.

34.3.5.1.8 Product discharge

At the conclusion of the cool-down period, the product is ready for discharge. Depending on the type of pan used, the product will be discharged through a "trap-door" on the flat of the pan or through the front

pan door (either by manual scooping, or using one of the many emptying scoop devices supplied by the equipment vendor). In some systems, the pan may be rotated backward or counter-clockwise to aid in discharging the product, in which case the baffles are designed to direct the product out the front door. Discharge of the entire batch usually will be achieved in 2–10 minutes, depending on the pan size and type of discharge method employed. For most cases, the coated product should not be allowed to drop more than 60–90 cm, otherwise tablet breakage may be experienced.

34.3.6 Understanding process thermodynamics

34.3.6.1 Adequate evaporative rate

The theoretical evaporative rate for aqueous film coating can be determined using the following equation:

$$\text{Rate} = \frac{\text{CMH} \times C_P \times \ell \times [(T_{IN} - T_{OUT}) - 0.10(T_{IN} - T_{OUT})]}{\text{LHV}}$$

where:

Rate = Evaporative rate of water per hour (kg)
CMH = Actual process airflow in cubic meters per hour
C_P = Specific heat capacity of the air (kcal/kg °C)
ℓ = Density of the air stream (kg/m^3)
T_{IN} = Inlet process temperature (°C)
T_{OUT} = Exhaust process temperature (°C)
H_L = Heat loss of the coating system (%)
LHV = Latent heat of vaporization (kcal/kg °C)

This equation will determine the approximate quantity of water (in kg/hour) that can be evaporated in a coating-pan system. However, this may not be the optimal rate necessary to achieve a high-quality coated product. The quality level, as stated previously, depends on many other factors. If all of these other critical factors are examined and optimized, this equation can be used as a tool to determine the approximate spray rate.

Example:
Assumptions:
Ambient air conditions = 21.1°C/50% RH
Air flow = 2166 m^3/hour
Specific heat capacity of air = 0.241 (kcal/kg °C)
Density of the air stream = 1.015 (kg/m^3)
Inlet air temperature = 70°C (167°F)
Exhaust air temperature = 40.5°C (105°F)
Coating solution solids = 12%
Number of spray guns = 4
Heat loss = 10%
Latent heat of vaporization = 577 (kcal/kg)

$$\text{Rate} = \frac{2166 \times 0.241 \times 1.015 \times [(70.0 - 40.5^\circ\text{C}) - 0.10(70.0 - 40.5^\circ\text{C})]}{577}$$

$$= 28.51 \text{ kg/hour or } 475 \text{ g/minute of water evaporated}$$

For a 12% w/w solids coating liquid, this equals 540 g/minute per total or ~135 g/minute per gun.

34.3.6.2 Process air

34.3.6.2.1 Volume

In all coating pans, elevated temperatures are used in conjunction with airflow to convert the solvent in the coating solution into a vapor and carry it away from the tablets being coated. In general, the amount of water vapor that can be removed is directly proportional to the air volume that passes through the coating pan and is limited by the saturation of the air stream. Therefore, if maximum evaporative capacity is the rate limiting factor with respect to the spray rate, it would be advantageous to use as high an airflow as possible without creating spray turbulence or air leakage. Manufacturers of the coating pans will usually specify the maximum airflow that can be used while still maintaining an acceptable level of turbulence. As the process air stream is heated, its capacity to hold water vapor increases. Again, if maximum evaporative capacity is the rate limiting factor, the inlet air temperature should be as high as possible. However, spray rates (per spray gun) are frequently limited not by evaporative capacities, but rather by the diminishing quality of the spray as the spray rate is increased. The total spray rate is usually limited by the number of spray guns, but more accurately, the size of the spray zone available. Therefore, the airflow capacity between pans of different sizes should be in direct proportion to the spray zone. For example, if a coating process is developed in a pan with two spray guns (each gun possessing a 20-cm spray pattern) and a process airflow of 1000 CMH (cubic meters per hour); scale-up to a larger pan with four spray guns should have an airflow of approximately 2000 CMH. Using an airflow in direct proportion to the spray rate will also allow a closer correlation between the process temperatures in different size pans.

34.3.6.2.2 Humidity

The process air used in a coating pan is either conditioned or unconditioned. In either case, there are day-to-day variations in the moisture present in the air stream. Humidity in the inlet air stream is usually determined by the dew-point temperature. Dew-point temperature provides a direct indication of humidity

in the air. Relative humidity can be used; however, the temperature of the air must also be known to determine the level of moisture. In the case of conditioned air, variations are less (especially when the process has the capability to dehumidify and rehumidify the air stream). With unconditioned air, these variations are great and can cause coating problems. If the air is unconditioned, the spray rate must be selected such that in the most humid condition that might occur, the film coating would still dry at a rate that would not result in overwetting or stability problems with the product. Some level of spray-drying would occur since most batches would be coated at lower humidities. The severity of the spray-drying would be directly related to the variation in the ambient humidity. Therefore, the growing trend is to dehumidify the process air stream. Dehumidification not only provides the coating process with more consistent coating conditions, but also provides for greater evaporative capacity. This allows the system to evaporate more moisture at a given airflow and temperature than an airflow of a higher humidity. With the advent of sustained release coating, some companies have opted for a combination of dehumidification/humidification. Such a system allows the process to be conducted at a consistent inlet dew-point temperature, regardless of the ambient conditions. A system that conditions the inlet air stream to a consistent dew point allows the coating process to operate at a reproducible drying rate regardless of fluctuations in ambient conditions.

34.3.6.2.3 Temperatures

As stated previously, the coating process can be successfully controlled by the inlet, exhaust, or product temperatures. Process control based on inlet air temperature control is most common. Control based on exhaust or product temperatures will often react slower because of the heat sink effect of the tablet bed (such control will also require fluctuations in inlet temperature, spray rate, and process air volume to occur, with spray rate fluctuations being the least desirable). With inlet temperature control, the exhaust temperature will drop slightly after the spray is started. This temperature drop is due to evaporative cooling. This will not occur with exhaust temperature control because the inlet temperature will be controlled to maintain the desired set-point. In any event, control via any of these methods will yield the same approximate temperatures. All of these temperatures are equally important because the moisture from the spray droplets is dried by both convection (due to the process air) and conduction (due to the product temperature). The desired exhaust temperature is dependent on several factors, such as:

1. Coating solution characteristics

   The tackier the coating as it dries, the higher the exhaust or product temperatures must be to prevent overwetting defects during the coating process. If the film polymer is thermoplastic, the product must be kept below the temperature at which the polymer begins to soften to prevent the tablets from blocking or sticking together.
2. Product temperature limits

   If a product exhibits instability problems at elevated temperatures, the product must be kept below these limits. This is most critical during preheating of the tablet bed since evaporative cooling of the bed is not occurring and the product may not be uniformly heated if the pan is not being continuously rotated. The product temperature will begin to rise immediately after the spray is stopped due to the loss of the evaporative cooling effect. Therefore, it may be necessary to begin a cool-down cycle immediately after the spray cycle.

   The inlet temperature required to achieve the desired exhaust temperature will be affected by the spray rate, the percentage of solids in the coating solution, the heat loss across the pan, and the condition of the drying air.
3. Pan speed

   To effectively optimize film-coating quality, the tablets must be mixed such that each tablet has the same probability of being in the spray zone for an equal duration of time. Therefore, it is essential that the product be examined to ensure that mixing is uniform. Mixing problems that occur include: sliding of tablets (usually seen with large capsules or oval-shaped tablets), “dead-spots” or sluggish product movement (this can be an extremely serious problem if it occurs while the tablets are in the spray zone), or product thrown into the spray zone by the mixing baffles.

   If product movement is not uniform, the first course of action should be to evaluate the product flow at different pan speeds. The pan speed selected should be the lowest speed that produces a rapid and continuous product flow through the spray zone. This will allow for the uniform application of a film coating while subjecting the tablets to a minimal amount of attrition. In general, if tablet friability is less than 0.1%, tablet attrition will not be a problem. A smaller tablet can be slightly softer since these tablets produce a less abusive tumbling

action. Product flow can be evaluated either subjectively or through mixing studies using tablets of various colors. Tablets of different colors can be placed in different zones (eg, front, middle, and back) of the pan. Tablet samples are taken at set time intervals to determine the length of time required to achieve a homogeneous mixture. A more sophisticated means of evaluation includes the use of radioactively marked tablets and a counter mounted on the spray bar to record the number of passes through the spray zone per unit time. The product speed is generally assumed to be traveling at the same linear velocity as the inside circumference of the coating pan.

In a continuous-coating pan system the product flow is from the charge port to the discharge port. Depending on the coating level, the product may pass through multiple drums (these drums may be located end-to-end or stacked horizontally). One continuous-coating pan system currently available contains five spray guns for a 1.5 m coating-drum length.

34.3.7 Understanding spray dynamics

34.3.7.1 Spray rate

The selection of the proper spray rate is dependent on more than just the thermodynamic considerations. If this were not so, the ultimate coating system would offer unlimited airflow and temperature. The spray rate (per spray gun) is also dependent on the ability of each spray gun to produce a consistent droplet size distribution. It has been shown that the droplet size will increase as the spray rate is increased (unless a compensating change is made in the atomizing air). Other factors that must be considered when determining the spray rate are:

- Solution viscosity—As solution viscosity is increased, the spray gun's ability to produce an acceptable droplet size distribution is diminished. Therefore, viscosity will limit the maximum spray rate that can be used and still produce acceptable film quality.
- Spray pattern width—If the spray pattern width is set-up properly, the spray pattern will essentially be the same as the spray gun spacing. The typical gun-to-gun spacing is 12–20 cm. At spacing greater than 20 cm, the uniformity of the spray across the pattern begins to deteriorate. Spacing of less than 12 cm is an ineffective use of spray guns and tends to add more expense for added spray guns, solution lines, and pumps. There is a limit to how much spray can be applied to a tablet per pass through the coating zone before the tablet begins to exhibit coating defects. Therefore, the wider the spray pattern width (without overlapping adjacent spray patterns), the greater the spray rate per gun.
- Product movement—The more consistent and fast the product flow through the spray zone, the higher the spray rate that can be delivered and still achieve an acceptable level of film-coating uniformity. Product movement is often dictated by the pan speed, baffle design, tablet size and shape, and tablet robustness (particularly friability).

34.3.7.2 Droplet size distribution

One of the most critical aspects of coating concerns is the manner in which solution is applied onto the tablet. The spray droplets can be almost any size if the size distribution is sufficiently narrow. If the system is set up to operate such that smaller droplets dry properly, then larger droplets will stay wet and picking or twinning will occur. If the system is set up to dry larger droplets properly, small droplets will be dry and not spread properly on the tablet, thus causing "orange-peeling" and a dull, rough film appearance will result. Most spray guns used in the industry today have a limited range over which the spray can be varied and still retain a uniform distribution. Examples of some typical droplet size distributions are shown in Fig. 34.30. These droplet size distributions were produced using a coating solution with a viscosity of 130 cP. Ideally, a spray gun that produces a narrow droplet size distribution should be used. This distribution can vary with changes in solution viscosity, spray rate, or a change in solution solids; therefore, the manufacturers of the spray guns will generally not provide information concerning these parameters. A simple method of examining this distribution can be performed by quickly passing a sheet of paper through the spray and subjectively

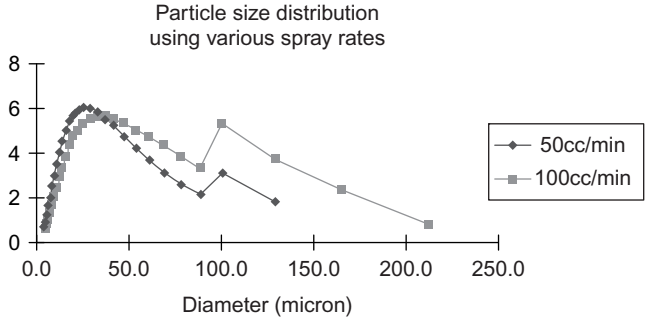


FIGURE 34.30 Examples of droplet size distributions obtained at different spray application rates.

analyzing the droplet size. A typical droplet distribution usually ranges from 5 to 250 μm. In general, if droplets vary in size from 5 to 1600 μm, it will be difficult to optimize the film-coating quality.

34.3.7.3 Coating zone/pattern

As mentioned previously, the larger the spray zone per spray gun, the higher the maximum spray rate that can be used per spray gun. Therefore, if the gun spacing or the spray pattern is reduced, the spray rate should be reduced proportionally. Applying more spray per unit area of the tablet bed beyond a certain point will change the film-coating appearance. If the spray rate is increased above this maximum, either overwetting or spray drying will occur, depending on whether the process is run dry or wet.

A recent trend in coating is to achieve greater production through the use of elongated or "stretch" coating pans. Lengthening the coating pan not only increases its capacity but, more importantly, the spray zone. This allows more spray guns to be used thereby increasing the overall spray rate. An increase in pan volume achieved by increasing bed depth with no increase in spray zone does nothing to allow an increase in overall spray rate. Therefore, one means of scaling up the process time for coating can be achieved by using the following calculations:

$$\text{Spray time (large pan)} = \text{Spray time (small pan)} \times \frac{\text{Batch size (large pan)}}{\text{Batch size (small pan)}} \times \frac{\text{Spray zone (small pan)}}{\text{Spray zone (large pan)}}$$

These calculations assume that the coating zones for each of the pans are used efficiently.

A number of factors have an effect on the spray pattern width. As the pattern air volume for the spray gun is increased, the spray pattern widens. An increase in the atomization air will cause a reduction in the pattern width. This is logical because the increased volume of atomization air would make it more difficult to flatten or widen the pattern. Lastly, as the spray rate is increased, the spray pattern is widened. Therefore, if the atomization air volume or spray rates are adjusted, it then is important to reexamine the spray pattern width to ensure that the spray zone is effectively utilized. Balancing the atomizing and pattern air correctly is critical to maximizing the size of the spray zone while avoiding undesirable effects, such as creating a dumbbell-shaped spray pattern.

34.3.7.4 Coating analysis

These are some characteristics of a typical film coating process:

Product specification:	Tablet	8.0 mm diameter
	Weight	250.0 mg/tablet
	Density	0.73 g/mL
Coating solution specifications:	Solids	12% w/w
	Density	1.0 g/mL
	Quantity	62.5 kg
Equipment specifications:	Pan diameter	170 cm
	Pan volume	550 L
	Pan speed	5 RPM
	Four spray guns	
Coating conditions:	Spray tate	125 g/min per gun
	Total spray rate	500 g/min
	Spray time	125 min

- Assumptions:
  - Spray zone:

Per spray gun = Spray width × Spray pattern length = 7.5 cm × 20.0 cm = 150 cm^2

Total spray zone = Number of spray guns × Spray zone per gun = 4 × 150 cm^2 = 600 cm^2

Tablets per coating zone = Total spray zone × Number of tablets/cm^2

Assume 3.5 tablets/cm^2 in spray zone = 600 cm^2 × 3.5 tablets/cm^2 = 2100 tablets/coating zone

Total spray length = Spray pattern width × Number of spray guns = 20 cm × 4 = 80 cm

Tablet bed velocity (same as the peripheral pan velocity) = Pan circumference × Pan speed
= 3.14 × 170 cm × 5.0 RPM = 2669 cm/minute

- Tablets per minute in the spray zone:

$$\begin{aligned}\text{Coating zone/minute} &= \text{Tablet bed velocity} \times \text{Total spray length}\\ &= 2669\ \text{cm/minute} \times 80\ \text{cm} = 213,520\ \text{cm/minute}\\ \text{Tablets/minute in the spray zone} &= \text{coating zone/minute} \times \text{Number of tablets/in}^2\\ &= 213,520\ \text{cm/minute} \times 3.5\ \text{tablets/cm}^2\\ &= 747,320\ \text{tablets/minute}\end{aligned}$$

- Coating solution per pass through the spray zone:

$$\text{Coating per pass through the spray zone} = \frac{\text{Total spray rate}}{\text{Tablets per minute/spray zone}}$$

$$\begin{aligned}\text{Coating per pass through the spray zone} &= \frac{500\ \text{g/minute}}{747,320\ \text{tablets/minute}}\\ &= 0.669\ \text{mg/tablet}\ (0.080\ \text{mg solids})\end{aligned}$$

- Passes through the spray zone:

$$\text{Total passes through the spray zone} = \frac{\text{Tablets per minute/spray zone}}{\text{Total number of tablets}} \times \text{Spray time}$$

$$\text{Total passes through the spray zone} = \frac{747,320\ \text{tablets/minute}}{1,700,000\ \text{total tablets}} \times 125\ \text{minutes} = 55.0\ \text{passes}$$

- Weight gain:

$$\text{Weight gain/tablet} = \text{Coating per pass through the spray zone} \times \text{Number of passes}$$
$$\text{Weight gain/tablet} = 0.080\ \text{mg solids/pass} \times 55.0\ \text{passes} = 4.40\ \text{mg}$$

$$\text{Total weight gain} \approx \frac{\text{Weight gain/tablet}}{\text{Tablet weight}}$$

$$\text{Total weight gain} \approx \frac{4.40\ \text{mg}}{225\ \text{mg/tablet}} \approx 2.0\%$$

In this analysis, each tablet would be within the spray zone approximately 55 times over the course of a 125-minute coating trial. In actuality, many of these tablets would be in the spray zone either more or less frequently, depending on the uniformity of the product movement. In addition, failure to optimize the gun-to-gun delivery or spray distribution will have an adverse effect on the coating uniformity. Therefore, this factor should be examined to ensure a consistently high-quality film coating.

34.3.8 Controlling coating processes—critical factors

Once all formulation requirements have been fulfilled, these factors can be examined to optimize the coating quality and system performance, and can generally be grouped into three areas: uniformity of spray application, uniformity of product movement, and achieving an adequate evaporative rate. These factors must all be examined and optimized in order to optimize the coating process.

34.3.8.1 Uniformity of the spray application

34.3.8.1.1 Spray gun design

There are a variety of spray gun manufacturers. Some of the more common suppliers include Binks, Freund, Schlick, and Spraying Systems. The most common type of gun used is air atomizing or pneumatic. This type of gun allows for the use of variable spray rates. In the past, hydraulic gun systems were used; however, they did not allow flexibility in spray rates. The solution nozzle orifice must be sized to match the desired spray rate. In order to change the spray rate, a change in the fluid nozzle may be necessary. In addition, the use of smaller nozzle orifices can

increase susceptibility to plugging of the gun tip. With some types of spray guns, the air cap and spray-nozzle configuration produce a predetermined ratio of atomization to pattern air at a given supply pressure. In other guns there are separate controls for adjusting the volumes of the atomization and pattern air. This allows the separate adjustment of either atomization or pattern without having to change the air cap and/or the solution nozzle. For example, the width of the spray pattern can be changed without changing the spray droplet size. The atomization air breaks the solution stream into a fine droplet size while the pattern air serves to flatten the spray into a fan-shaped pattern. The volume of atomization controls the mean droplet size of the spray. The pattern air volume controls the overall width of the spray.

34.3.8.1.2 Number of spray guns

The number of spray guns needs to be adequate to provide uniform coverage of the entire product bed. To maximize the uniformity and application of the coating, the spray zone should cover from the front edge to the back edge of the tablet bed. Adding more spray guns will not automatically guarantee that the overall spray rate can be increased. The use of additional guns is only justified if the existing number of guns is insufficient to cover the tablet bed from the front to back of the pan. The objective is to produce a uniform "curtain" of spray that the tablets pass through. Spray guns are usually capable of developing pattern widths of 12–20 cm without adversely affecting the spray droplet distribution. At greater pattern widths, the volume of pattern air needed to fan out the spray can lead to a distortion of the droplet size distribution due to recombination of spray droplets. The spray guns should be set up such that adjacent spray patterns are as wide as possible without overlapping. Overlapping of spray patterns can lead to localized overwetting of the tablet bed.

34.3.8.1.3 Uniform gun-to-gun solution delivery

It would seem obvious that to achieve uniformity of solution application, all spray guns must be set up to deliver the same quantity of coating solution. A recent trend in coating systems is the use of a single-pump manifold for multiple spray guns. For these types of systems, it is mandatory that calibration be performed on a regular basis to ensure that all spray guns are delivering the same quantity of solution. Furthermore, the calibration must be performed with the coating solution to be used. Calibration with water will not be satisfactory because its viscosity is much lower than that of the coating liquid and it will not be as sensitive to differences in pressure drop between the spray guns. Calibration is also recommended when using peristaltic pumps because the tubing is subject to fatigue. Calibration is usually accomplished by adjusting a knob that controls the restriction of the spray nozzle by the spray needle.

34.3.8.1.4 Atomization air volume/droplet size

As stated previously, the atomization air volume can be adjusted to control the mean droplet size of the spray. An increase in the atomization air volume can reduce the mean or average droplet size. An increase in either spray rate or solution viscosity will cause an increase in not only the droplet size distribution, but also the mean droplet size. Therefore, the droplet size should be evaluated at the exact spray rate that will be used for the coating trial. The quality of the spray in terms of droplet size and distribution should be evaluated at several different settings.

34.3.8.1.5 Spray gun angle

Ideally, the spray gun should be directed at the middle (midway between the leading and trailing edges of the tablet bed) and at a 90 degree angle to the moving tablet bed. If the spray guns are directed higher toward the leading edge of the tablet bed, it is possible that spray could be applied to the pan or onto mixing baffles as they begin to emerge from the tablet bed. Conversely, if the spray guns are directed too low on the tablet bed, spray applied to the tablets may not have sufficient time to dry. This could result in the transfer of film from the tablets to the pan surface. If the spray guns are not directed at a 90 degree angle to the tablet bed then, the spray as it exits the solution nozzle has a tendency to build up on the wings of the air cap.

34.3.8.2 Uniformity of product movement

34.3.8.2.1 Pan speed

This area was covered in an earlier section (see Process Parameters-Pan Speed). Product movement must be uniform if a uniform application of the coating solution is to be achieved and the minimum pan speed necessary to achieve this objective is recommended. Once the pan speed has been determined, it can be scaled up by duplicating the peripheral edge speed. This is done by taking the ratio of the small pan to large pan diameter multiplied by the small pan speed.

One final note is that the product movement should be continually evaluated and adjusted as necessary during the coating cycle. Frequently, the product movement will change as a coating is applied to the tablet surface and the degree of slip of the tablets increases.

34.3.8.2.2 Tablet size and shape

Different tablet sizes and shapes will exhibit very different flow characteristics. In general, smaller tablets will flow better than larger tablets. Longer, less round shapes, such as capsule and oval shaped tablets will tend to slide and flow more poorly (more sliding) than other shapes. If the tablet shape or size is changed, the product flow properties must be reexamined. In addition, if the size is too small, it will require a fine mesh to cover the perforations in the pan to prevent the exhaust of product. One other consideration is that if the product is fine (~1 mm or less) and the airflow is exhausted down through the product bed, then there will be a considerable pressure drop across the bed. This pressure drop will cause a reduction in the process airflow.

34.3.8.2.3 Baffle type/size/number

The primary function of the mixing baffles is to transfer the product between the front and back of the coating pan. A variety of different baffle shapes and sizes exist. Coating pans are usually equipped with a standard baffle design that works well for the majority of different products. However, it may be necessary to use a different baffle design for unusual shapes and sizes. Coating pans can also be fitted with anti-slide bars. These bars are positioned on the flat of the coating pan and perpendicular to the product flow. They are used to prevent tablets with large flat surfaces from sliding inside the pan. With most coating pans, it is necessary to use a reduced baffle size when working with smaller batch sizes.

Standard-size baffles used with a small batch will result in sluggish product movement and/or excessive variation in the spray gun-to-product distance. As the baffle passes through the product, it will temporarily carry a portion of the tablets up out of the tablet bed, thus causing a brief increase in the gun-to-bed distance. As these tablets cascade off the baffle, the tablet bed height rises and the gun-to-bed distance decreases. A minimum variation in the gun-to-bed distance is desired so that the spray always travels a consistent distance and spray droplets striking the tablets have a constant moisture level. The typical variation in gun-to-bed distance is from 2.5 to 5.0 cm. A reduced or small-batch baffle is generally recommended whenever the batch size is less than 75% of the rated pan brim volume.

34.3.8.2.4 Batch size

As previously mentioned, the batch size/baffle combinations are critical to obtain an acceptable product movement. An acceptable batch size range for film coating is usually 50–95% of the rated brim volume. By using only 95% instead of 100% of the rated volume, one can eliminate the spillage from the pan mouth ring during coating. Occasionally, what appears to be an acceptable pan load initially may turn out to be excessive for those products that exhibit a change in product movement as a film is applied or with coatings of an extremely high weight gain. A problem in working with smaller batches (~50%) in a perforated coating pan is that, unless part of the exhaust plenum is blocked off, the process air will preferentially pass around the tablet bed due to less restriction or pressure drop. Batches of this size potentially can be coated successfully, however, the drying efficiency will be reduced.

34.3.8.3 Adequate evaporative capacity

34.3.8.3.1 Process air volume

The process air stream should be adjusted to the maximum volume that yields a laminar, nonturbulent air flow. A turbulent airflow will distort the spray patterns and lead to lower coating efficiency due to spray drying of the coating. It is important to periodically inspect the coating pan to ensure that all process air passes through the tablet bed. Any air that passes around the product will result in a reduced evaporative efficiency. The inlet air volume is more important than the exhaust air volume as an indicator of the evaporative capacity because it is this air that passes through the tablet bed and vaporizes the water from the coating and then conveys it away from the tablets. The exhaust air volume will typically be slightly greater than that of the inlet due to the addition of atomizing air from the spray guns, also due to slight leakage that may exist, and by the addition of water vapor to the exhaust air stream. If there is a large difference between the inlet and exhaust air volumes due to leakage, this can be a problem, because this air can artificially depress the exhaust temperature.

34.3.8.3.2 Spray rate

The initial spray rate that is selected may be based on previous coating trials that have yielded successful results. This is an acceptable approach if the coating solution and product being coated are very similar. However, it is important to remember that changes in these factors can drastically affect the spray rate and other parameters selected. Another method for determining an initial spray rate is to evaluate the quality of the droplet size distribution. Spray rates for aqueous film coating vary from 6 to 30 g/minute for a small 2.0 L pan to 80–250 g/minute per gun in a large production-scale pan. The key factors that limit the maximum spray rate per gun are the viscosity of the coating solution, the type of spray gun used,

and ultimately the level of film quality that a Pharmaceutical company deems acceptable.

34.3.8.3.3 Spray gun-to-tablet-bed distance

For small-scale coating systems, the gun-to-bed distance can be as little as 2.5–5.0 cm. The typical gun-to-bed distance is 20–25 cm for a production-sized coating pan. This distance usually provides an economical tradeoff between the cost of the number of guns needed to adequately cover the spray zone and the desired quality of spray. If the gun-to-bed distance is less than 20 cm, then either the spray rate must be reduced or the inlet temperature and product temperature increased to compensate for the shortened evaporation time. If this distance is greater than 20 cm, the inlet process temperature should be reduced otherwise more spray drying is likely to occur.

34.3.8.3.4 Product/exhaust temperature

Typically, aqueous film-coating processes are set up to achieve an exhaust temperature of 38–44°C. Based on the desired spray rate, an inlet temperature is determined that allows an exhaust temperature in the target range to be maintained. Coatings that develop greater tackiness will require exhaust temperatures to be adjusted to prevent overwetting defects, or excessive heat when working with thermoplastic coatings that become tacky when getting too hot (for these coatings, a typical exhaust temperature range might be 30–38°C).

The exhaust temperature is slightly lower (~1–5°C) than that of the product bed due to heat loss between the measurement points. Usually, the greater the distance between the exhaust and product temperature probes, the greater the differential. The only time when these temperatures will vary more is during preheating and cool-down. A product temperature probe will display an average of the entire tablet bed. The product temperature can be determined through the use of either a probe that extends into the tablet bed or through the use of an infrared temperature probe directed at the zone just above the spray zone. Tablets in the spray zone are at a slightly lower temperature due to evaporative cooling and are not representative of the average product temperature.

34.3.8.3.5 Dew-point temperature

The dew-point temperature is directly related to the moisture contained in the process airstream. Dew-point temperatures can be measured using either a capacitance or a chilled mirror-type dew-point sensor. To accurately reproduce the drying rate from trial to trial, it is recommended to maintain the dew point within a controlled range. The more critical the coating (ie, sustained release coatings), the tighter the range. High dew-point temperatures can be reduced through the use of a dehumidification system employing either chilled water coils or desiccant dehumidification. Chilled water systems are usually specified to control the dew point at 10–12°C (50–53°F) or an absolute moisture content of 7.5–8.5 g/kg of air. With a desiccant dehumidification system, the dew point can routinely be controlled to a temperature as low as −6°C (21°F). Low dew-point temperatures can be adjusted by humidifying the air via the injection of clean steam into the process air. If no attempt is made to limit the variation of the inlet dew point, then fluctuations in ambient air conditions can lead to reduced coating efficiency (spray drying), longer processing times, or film defects due to overwetting.

34.3.9 Scale-up

34.3.9.1 Batch size

The most accurate method of determining batch size is to load the coating pan to within 2.5–5.0 cm of the pan opening and then rotate the pan at the desired rpm to ensure that the pan is not overfilled. However, there are two different methods for approximating the batch size. The first method is to multiply the rated brim volume times 95%, and then multiply the resultant volume by the bulk density of the product. A second method for determining the pan load size is to multiply a known ratio of batch size to pan volume for a small-scale pan times the volume of the pan being scaled to. For example:

$$\text{Large pan batch size} \approx \frac{\text{Small pan batch size}}{\text{Small pan volume}} \times \text{Large pan volume}$$

$$\text{Large pan batch size} \approx \frac{65.0\ \text{kg}}{90.0\ \text{L}} \times 550.0\ \text{L} \approx 397.0\ \text{kg}$$

However, on the production scale, pan fill weights are often determined by compression batch weight rather than desired pan fill capacity, so coating on this scale may use pan loadings that are less than ideal.

34.3.9.2 Pan speed (angular pan velocity)

When scaling up a coating process, it is critical that the tablet speed through the spray zone in the larger pan is comparable to that used in the smaller pan. In other words, the pan angular velocity must be the same for both coating pans. Pan velocity can be duplicated by multiplying a ratio of the small pan diameter

to large pan diameter times the pan speed used for the smaller coating pan. For example:

$$\text{Pan speed for large pan} = \frac{\text{Small pan diameter}}{\text{Large pan diameter}} \times \text{Pan speed for small pan}$$

$$\text{Pan speed for large pan} = \frac{100\ \text{cm pan}}{170\ \text{cm pan}} \times 9\ \text{rpm} = 5.3\ \text{rpm}$$

This equation will yield a close estimate of the pan speed. However, subtle differences in the baffle design between coating pans may require a slight adjustment from this calculated pan speed.

34.3.9.3 Available coating zone

To scale up the coating process with any degree of confidence, the spray rate must be determined using the same gun-to-bed distance used in the larger coating pan. Any change in the gun-to-bed distance will change the drying time for the spray droplets and thus alter the quality of the coating. The same spray rate used in the small-scale pan can also be used in the production-scale pan, assuming that the same spray gun spacing and spray pattern widths are used. Typically, the use of more spray guns without increasing the size of the overall spray zone will not allow an increase in either the total or the per gun sprays. Increasing the total number of spray guns will only be of value if the existing spray zone is inadequately covered with fewer spray guns. If the spray pattern width used in the larger coating pan is narrower than that used in the small pan then the spray rate per gun should be reduced in proportion to the reduction of the spray zone width. This will allow the same density of film coating to be applied per unit area of the tablet bed surface. Otherwise, the quantity of coating applied on the tablets per pass through the spray zone would increase. This would most likely change the quality of the coating and could lead to overwetting or logo bridging defects. Here is an example of scale-up using a ratio of total spray zone utilized:

	Pan #1	Pan #2
Pan volumes	90.0 L	850.0 L
Batch size	60.0 kg	566.0 kg
Coating solution	2 kg	113 kg
Number of spray guns	2	10
Spray pattern width (per gun)	20 cm	15 cm
Spray rate (per gun)	150 g/min	To be determined
Spray rate (total)	300 g/min	To be determined
Spray time	40 min	To be determined

$$\text{Spray rate for large pan} = \frac{\text{Pattern width for pan \#2}}{\text{Pattern width for pan \#1}} \times \text{Spray rate (per gun) for pan \#1}$$

$$\text{Spray rate for large pan} = \frac{15\ \text{cm spray pattern}}{20\ \text{cm spray pattern}} \times 150\ \text{g/minute per gun} = 113\ \text{g/minute per gun}$$

$$\text{Total spray rate for pan \#2} = \text{per gun spray rate (pan \#2)} \times \text{number of spray guns (pan \#2)}$$
$$\text{Total spray rate for pan \#2} = 113\ \text{g/minute per gun} \times 10\ \text{spray guns} = 1,130\ \text{g/minute/total}$$

$$\text{Spray time} \approx \frac{\text{Quantity of solution to apply}}{\text{Total spray rate}}$$

$$\text{Spray time} \approx \frac{113\ \text{kg}}{1.33\ \text{kg/minute}} \approx 85\ \text{minute}$$

34.3.9.4 Spray-rate-to-pan-speed ratio

One factor that is quite commonly overlooked is the ratio of the spray rate to the pan speed. The ratio of spray rate to pan speed has serious implications in both the amount of film coating applied onto the individual tablets per pass through the spray zone and on the overall uniformity of the film coating itself. Therefore, with any increase in the spray rate, one should evaluate the need for an increase in the pan speed. In some instances, the tablets may be able to withstand a greater application per pass through the spray zone without any adverse effects on film quality or the uniformity of the coating. Whenever the spray rate and pan speed are both increased, the tablets should be evaluated for signs of overwetting or increased tablet attrition. One problem that may result if the pan speed and spray rates are too high relative to the evaporative rate is that wet film coating from the tablets may be transferred to the pan surface.

34.3.9.5 Airflow-to-spray ratio

When scaling-up the film coating process, it is recommended that the airflow used in the larger coating pan be proportional to the increase in the spray rate. If the airflow is increased in the same ratio as the increase in spray rate and if the spray rate per unit area of the bed surface is the same, the same inlet and exhaust temperatures can be maintained. The inlet temperature must be increased to maintain the same evaporative rate if the airflow is not increased in the same proportion as the spray rate. If the spray rate per gun and the gun-to-bed distance are the same for both pans, the airflow can be scaled up in direct proportion to the increase in spray guns. For example:

	Pan #1	Pan #2
Spray rate (per gun)	125 g/min	125 g/min
Number of spray guns	2	4
Total spray rate	250 g/min	500 g/min
Inlet airflow	1000 CMH	To be determined

$$\text{Airflow for pan \#2} = \frac{\text{Total spray rate for pan \#2}}{\text{Total spray rate for pan \#1}} \times \text{Airflow for pan \#1}$$

$$\text{Airflow for pan \#2} = \frac{500\ \text{g/minute}}{250\ \text{g/minute}} \times 1000\ \text{CMH} = 2000\ \text{CMH}$$

34.4 TROUBLESHOOTING

34.4.1 Introduction to troubleshooting

Troubleshooting is basically a "reactive" process because it deals with something that has already gone wrong. A coating problem will usually manifest in one or more of these ways:

- Those affecting visual coated-product quality,
- Those affecting coated-product functionality,
- Those affecting coated-product stability, and
- Those affecting processing efficiencies and costs.

When dealing with an existing, marketed product, the troubleshooting process is constrained by many regulatory issues. These problems are the most troublesome in that the proper corrective action may require refiling with the FDA. Therefore, other less suitable remedies may be taken, which may lessen the symptoms rather than correct the actual cause of the problem. Of course, application of a suitable quality by design (QbD) program, as mandated by regulatory agencies, should essentially help eliminate problems in film-coating processes, with the possible exception of those due to equipment failure.

Identifying appearance-related problems is relatively easy because visual feedback is immediate. The magnitude of the problem is also often immediate. Identifying nonappearance related problems (such as those associated with chemical stability or drug release) is more difficult because:

- The existence of the problem is often not readily apparent,
- Determination is often on the basis of some analytical procedure that evaluates only a small sample (relative to the batch size in question) of tablets, and
- Sampling, and the relevance of the samples selected to the characteristics of the whole batch, becomes a critical issue.

34.4.2 Up-front approaches to avoid troubleshooting issues

The best solution to "fixing problems" is to avoid them in the first place. One way in which many coating problems can be avoided is through the proper formulation of the product substrate and the film-coating solution. Before the coating process can be developed or scaled-up, the product must be evaluated to ensure that it meets the formulation requirements for film coating. Often, problems arise because the tablet core formulation is not sufficiently robust to withstand the rigors of film coating.

The core must be formulated such that minimal attrition occurs during the film-coating process. The deeper the product bed, the more abuse the tablets must withstand. For smaller tablets (less than 100 mg), 5–6 kp tablet-breaking force (hardness) may be sufficient. For medium-sized tablets, 12–16 kp may be sufficient. For larger capsule-shaped tablet, such as a 1 g capsule-shape, a hardness of more than 20 kp may be necessary. A better measure of a product's suitability for coating is friability. In general, a product with a friability of 0.1% or lower should be sufficient to avoid attritional problems during the coating process. Other core formulation issues that should be addressed prior to developing the coating process include:

- Product stability: All core ingredients must be stable at the temperatures required for the evaporative process (typically 40–45°C for aqueous-based coatings)
- Product shape: The use of large flat product surfaces must be avoided to prevent the tendency for "twinning" of product cores

- Logo design: The core logo should be designed such that attrition and bridging of the logo do not occur
- Resistance to dimensional changes (avoiding ingredients that undergo thermal expansion, or expand through excessive moisture absorption)
- Film adhesion: The core should be sufficiently porous to allow the film to adhere properly to the surface
- Chemical/functional robustness:
  - Role of amorphous, hydrophilic materials
  - Low melting-point ingredients
- Mixing potential: Small, more round products will tend to flow well. Large or longer tablet shapes will tend to exhibit "sluggish" movement.

Another issue that must be addressed before the coating process can be developed is the film-coating formulation, which must be extensively evaluated to ensure that it meets the requirements for intended coating process. Some of the issues that must be addressed include providing:

- Sufficient film mechanical strength to prevent cracking or edge wear;
- Sufficient plasticizer to prevent the formation of a brittle film;
- Sufficient pigment to mask the color of the tablet core;
- Appropriate solids level that allows a coating to be applied quickly to protect the core from attritional effects;
- Sufficient adhesion between the coating and the core surface;
- A solution viscosity that is low enough to allow the production of a uniformly atomized spray; and
- Sufficient coating suspension stability, including absence of growth of microorganisms and prevention of sedimentation of suspended coating solids.

34.5 CONSIDERATION OF PRODUCT SUBSTRATE

To effectively develop the coating process, the product must be evaluated to ensure that it meets the necessary criteria for a substrate.

34.5.1 Hardness/friability

The tablet core must be capable of withstanding the rigors of tumbling in the coating pan. In a larger diameter coating pan, the bed depth will be greater and therefore, subject the tablets to greater stress. So, an acceptable tablet hardness or friability for a small coating pan may not be sufficient for a larger pan. A tablet hardness tester is used to determine the edge-to-edge (diametral) tablet hardness. Typical hardness measurement units include: kiloponds (kp), Strong Cobb Units (SCU), and Newtons (N). The kilopond is defined as the force exerted by a kilogram mass upon its support in a gravitational field of $g = 9.80665$ m/second^2. One kilopond is equal to 9.807 N units or 1.4 SCU. Tablet hardness has traditionally been the measure of a tablet's suitability for coating. However, in many cases, the tablet may be of substantial hardness but still exhibit unacceptable capping tendencies or show excessive wear on the tablet edges or logo. Therefore, a better means of determining a tablet's ability to withstand tumbling is friability. This is usually determined by tumbling a certain number or weight of tablets for a set number of rotations (usually 100 revolutions) inside a cylinder. The tablets are weighed before and after tumbling and weight loss is expressed as the percent friability. A recent innovation to the friability test is to line the inside of the friability cylinder with a mesh screen. This has been shown to provide a better correlation between the friability test and actual suitability for subjecting to a coating process.

34.5.2 Weight variation

Tablet cores are usually produced to a particular weight range specification. However, sometimes these ranges are not sufficiently narrow. This usually occurs when the tablet granulation exhibits flow problems. A wide tablet weight variation will make it difficult or impossible to accurately determine the actual tablet weight gain due to the application of the film coating, since the weight variation in the uncoated cores can be greater than the weight of the film to be applied. In addition, wide variations in tablet weight can be accompanied by variations in tablet hardness.

34.5.3 Stability

The tablets must be stable under the conditions required for coating. The product must be able to withstand the temperature and humidity of the process airflow. Product temperature is significantly less than that of the inlet air during the coating process due to evaporative cooling. However, during the preheat phase, the product temperature may approach that of the inlet air if the tablets are not jogged frequently enough. Tablets should be able to handle the usual product temperature of 35–50°C. As mentioned previously, it may be necessary to use drier conditions during the initial phase of the coating process to

prevent stability problems with moisture-sensitive products, such as effervescent tablets.

34.5.4 Compatibility

The compatibility of the tablet core with the excipients in the film-coating solution must be verified. In some instances, certain actives have exhibited an interaction with the plasticizer in the coating solution.

34.5.5 Shape

If possible, certain tablet shapes should be avoided for film coating. Tablets with sharp edges may exhibit a greater tendency for edge wear. Cores with large, flat tablet faces may result in poor product movement due to sliding. Tablets with large flat surfaces (such as deep walled capsule shaped tablets) will also show a much greater tendency to exhibit twinning during coating. Adding a slight concavity (0.1−0.2 mm) to the tablet faces or tablet edges reduces the likelihood of agglomeration or twinning.

34.5.6 Logo design

Sharp corners or small islands on the tablet logo can lead to logo attrition problems. If the logo is too fine or contains too much detail, the film coating may bridge or cover the logo. A draft angle of 35 degree is recommended for film-coated tablets. Tooling manufacturers are usually aware of the tool design specifications for tablets to be film coated. Finally, placement of a logo on the crown of a tablet may lead to more crown erosion.

34.5.7 Core porosity

The tablet core must be formulated so that there is good adhesion between the film coating and the tablet surface. If the core porosity is low, poor adhesion will result, and picking and/or peeling of the film will occur. Core porosity can be a problem with wax matrix tablets due to poor adhesion between the tablet surface and film-coating droplets. To remedy this, more adhesive film polymers (ie, hydroxypropylcellulose, copovidone, or polyvinyl alcohol) may be used.

34.5.8 Disintegration/dissolution

The disintegration of the tablet core must be sufficiently rapid when tested prior to the addition of the film coating. If the core does not dissolve quickly prior to coating, the addition of film coating may only provide a further delay. However, in most instances, a change in dissolution after film coating is often caused by exposure of the tablet core to the coating process conditions (such as heat) rather than as a result of a direct effect of the applied film coating.

34.6 COATING FORMULATION

34.6.1 Film mechanical strength

The film-coating formulation must be optimized such that the resultant film has adequate mechanical strength. The reasons for this are twofold; first, the film must be strong enough to protect the tablet from excessive attrition while tumbling during the coating process and, second, the applied coating must be sufficiently durable to resist the erosion of the film itself. A weak film will usually exhibit wear or erosion at the tablet edges, or may crack as a result of being subjected to mechanical stress (tumbling) or excessive heat. The addition of too many nonfilm-forming excipients (such as pigments and detackifying agents like talc) will decrease the strength of the film. Such reduced film strength may also occur if a drug is added to the coating solution.

34.6.2 Plasticizer level

The function of a plasticizer is to reduce the glass transition temperature (T_g) of the film or, in other words, to produce a film that is not brittle under normal process and storage conditions. If the film is inadequately plasticized, the film will be too brittle and more prone to cracking. If a film-coating formulation has an excessive level of plasticizer, the mechanical strength of the film will be reduced, and excessive tackiness may be observed.

34.6.3 Pigment level

If the coating solution contains insufficient pigment, it will be impossible to develop the desired color intensity. In addition, low pigment levels can make it difficult to minimize color variation due to poor opacity of the applied coating. On the other hand, excessively high pigment levels can, as stated previously, reduce the mechanical strength of the coating.

34.6.4 Film solution solids

A low solids level in the coating solution can not only needlessly increase the process time, but can also increase the time required to provide a protective film coating, thus resulting in increased tablet attrition. High solution-solids levels coupled with a low

addition level can make it difficult to achieve acceptable film-coating uniformity. High solids levels can also be a problem if they result in a high solution viscosity. If a high solids level does not create an excessively viscous solution, it provides an excellent opportunity to reduce the amount of coating solution required and, thus, the coating time. Aqueous polymer dispersions available on the market provide high solids solution at an extremely low viscosity.

34.6.5 Solution viscosity

Solution viscosity, as stated previously, is closely tied to the solution solids level. As solution viscosity is increased (above $\sim$200–250 cP) the droplet size distribution produced by the spray guns becomes increasingly wider due to an increase in larger droplets. Slight increases in viscosity can be compensated for by increasing the volume of atomization air. However, at higher viscosities (greater than $\sim$350 cP) it is difficult to eliminate all of the larger droplets. To compensate for the presence of larger droplets, the operator will usually increase the inlet temperature to prevent the larger droplets from overwetting the product. This will result in an increase in the amount of spray drying due to premature evaporation of water from the finer droplets. The net result is a lower coating efficiency and a rougher film surface.

34.6.6 Stability

The coating solution must be stable for the duration of the coating time. This includes both chemical and physical stability. The most typical instability problem encountered is settling-out of solution solids. This occurs with solutions that contain an excessive percentage of solids or when the coating has insufficient suspending capacity (due to low polymer content and/or low viscosity). Settling of solids can lead to blockage of the solution lines or the spray guns. For coating liquids where settling of insoluble components (such as pigments) is a problem, use of a recirculating spray system and/or using a suitable tank mixer can help.

34.6.7 Compatibility

This refers not only to the compatibility of the coating materials, but also to the compatibility between the film coating and the tablet core. For example, the addition of color concentrates that contain ethanol has led to precipitation of some of the polymer used in aqueous dispersions. As a remedy, a propylene glycol-based color was used. In another product there was an interaction between the plasticizer and the active ingredient used in the tablet core. It was necessary to reformulate the coating solution using an alternative plasticizer. Finally, when applying ionic coatings, such as enteric coatings, interaction with ionic materials in the core must be considered, as does potential interaction with core ingredients that are acid unstable (such as proton pump inhibitor APIs like omeprazole, lansoprazole, etc.)

34.6.8 Processing issues as they relate to troubleshooting

34.6.8.1 Equipment maintenance issues

Many coating problems can occur as a result of poor equipment maintenance, and these can be avoided through implementation of a routine equipment maintenance program. This program must involve the periodic calibration of all analog and digital instrumentation. Poor maintenance of process-monitoring equipment results in decisions being made on the basis of inaccurate information. The maintenance program should also involve the inspection of all wear items. The equipment manufacturer should be able to provide a listing of these parts with a recommendation for the frequency of inspection.

34.6.8.2 Process adjustment as a troubleshooting initiative

The dynamics associated with the spray application of the coating liquid represent perhaps one of the most underappreciated areas of the whole film-coating process.

Key issues to be aware of include:

- Fluid flow rate through the nozzle, interaction with the driving forces for effective atomization, and the ultimate size of droplets formed
- Droplet velocity
- The interaction of atomized droplets with the surrounding drying environment, and the relative state of the droplets as they arrive at the tablet surface

While the importance of many operating parameters (air flows, temperatures, spray rates, etc.) is relatively well understood, the ultimate impact of other parameters on the issue of troubleshooting is often overlooked, such as:

- Mixing effectiveness in the pan, and the impact of, for example, baffle design;
- Pan loading, and its impact on process efficiencies and product quality; and
- The potential change in tablet roll dynamics as coating is applied.

Another approach to troubleshooting involves the systematic evaluation of the film-coating process. Film-coating processing issues can be classified into these categories, namely achieving:

1. Good uniformity of the spray droplet size and application,
2. Effective uniformity of product movement through the spray zone, and
3. Proper solvent evaporation rates for the coating process.

There are a number of issues that must be addressed with regard to each of these objectives. This is a partial listing of some of the factors to be considered:

Issues that affect the uniformity of spray droplet size and application:

- Can an acceptable spray droplet size distribution be produced at the desired spray rate?
- Can an acceptable spray droplet size distribution be produced using the coating solution viscosity?
- Is there adequate compressed air volume for the atomization of the coating liquid?
- Are the number and type of spray guns suitable to provide uniform coverage of the product bed?
- Are the spray guns located at the proper spacing?
- Is the spray pattern width sufficient to maximize the size of the spray zone without overlap adjacent spray patterns?
- Is the spray-gun-to-product-bed distance correct?
- Is the solution delivery rate the same for all spray guns?

Issues that affect the uniformity of product movement:

- Is the batch volume sufficient for the pan being used?
- Is the correct pan speed being used?
- Does the pan contain the proper baffle type and are they in the correct position?
- Are tablets being "thrown" into the spray zone by baffles or baffle mounts?
- Is the product size and/or shape conducive to good product movement?

Issues that affect the solvent evaporation rate of the coating process:

- Is the proper process air volume being used?
- Is the incoming air stream controlled to a consistent dew point?
- Are the inlet, exhaust, and product temperatures correct for the tablets being coated?
- Is the proper spray rate being used?

It is important that the objectives be examined in the order listed to effectively troubleshoot the film-coating process. One should first deal with issues affecting uniformity of spray application, then issues that affect product movement, and lastly, issues dealing with the evaporative rate of the coating process.

For example, let's assume that the product is exhibiting signs of picking, which is an over-wetting defect. Let's also assume the root cause of this defect is due to poor product movement, which causes the product to stop in the spray zone. Now it might be possible to eliminate this problem by increasing the process temperatures (evaporative rate factor). However, the product movement through the spray zone would still be poor and coating uniformity would be less than optimal. A better remedy would be a corrective action that results in improved product movement (2nd Objective).

A second example might assume the same defect (picking), however in this case, the cause of the defect is due to nonuniformity of the solution delivery through the spray guns. One of the spray guns is delivering significantly more solution than the others. One could probably reduce the spray rate to eliminate the over-wetting problem; however this would not be the appropriate remedy. If this were done, then the coating uniformity would still be less than optimal and the corrective action would have lengthened the necessary coating-process time. A more appropriate remedy would be a corrective action which results in an improved spray application (1st Objective).

34.6.9 Troubleshooting: summary

Overall, problems often have more than one cause, and problem resolution is likely to be complex. Problem resolution may encompass minor changes that have little regulatory impact, or they may require intervention by making significant (from the regulatory standpoint) changes to either formulations or processes.

The simplest approach to problem resolution is problem avoidance.

34.6.10 Film-coating defects/ troubleshooting—summary

The information shown in Table 34.18 contains a brief definition of common tablet-coating defects together with typical causes and suggested remedies.

TABLE 34.18 Tablet Coating Defects and Suggested Corrective Actions

1. Overwetting/picking: This condition occurs when part of the film coating is pulled off one tablet and is deposited on another. If detected early in the process, it can be covered. If detected late in the process the coating will likely be unacceptable.

Possible cause	*Remedies*
Insufficient drying rate	Increase the inlet and exhaust temperatures
	Increase the process air volume
	Decrease the spray rate
Inadequate atomization	Increase the nozzle air pressure or the atomization air volume
	Decrease film coating viscosity
Poor product movement	Increase pan speed
	Switch to an alternative baffle design
	Adjust batch size (volume)
Poor distribution of spray	Check the uniformity of solution delivery through the spray guns (calibration)
Insufficient drying rate	Increase the inlet, exhaust, and product temperatures
	Increase process air volume
	Reduce the spray rate
Excessive variation in Process Air Humidity	Dehumidify and/or humidify process air to maintain a constant inlet air dew point

2. Twinning: Twinning is a form of overwetting whereby two or more of the tablet cores are stuck together.

Possible cause	*Remedies*
May be due to any of the possible causes for overwetting	
Poor tablet core design	Change tablet design to eliminate large flat surfaces

3. Orange peel**:** Appears as a roughened film due to spray drying. This condition relates to the level of evaporation that occurs as the spray droplets travel from the gun to the tablet bed. If there is excessive evaporation, the droplets do not have the ability to spread and form a smooth coating. A narrow droplet size distribution is important to ensure that the majority of droplets dry at the same rate.

Possible cause	*Remedies*
Excessive evaporative rate	Reduce the inlet and exhaust air temperatures
	Reduce the gun-to-bed distance
Excessive atomization of the spray	Reduce the nozzle air pressure or atomization air volume
Large droplet size variation	Reduce the solution viscosity
	Reduce the spray rate

4. Bridging: Bridging is a condition in which the film coating lifts up out of the tablet logo[10].

Possible cause	*Remedies*
Poor film adhesion	Reformulate the film coating solution to improve adhesion
	Reformulate the core formulation to increase porosity
Poor logo design	Redesign logo to incorporate shallower angles

5. Cracking: May occur due to internal stresses in the film[10].

Possible cause	*Remedies*
Brittle film coating	Increase the addition level of the plasticizer
	Use a different plasticizer
	Use a polymer with a greater mechanical strength
Poor film adhesion	Dilute the film-coating solution
	Reduce the quantity of insoluble film coating additives

(*Continued*)

TABLE 34.18 (Continued)

6. Poor coating uniformity: Poor coating uniformity can manifest itself in either a visible variation in color from tablet to tablet, or in the form of an unacceptable release profile for tablets.

Possible cause	*Remedies*
Insufficient coating	Apply a coating of 1.5−3.0% (for clear coating as little as 0.5%) weight gain to attain uniformity, requires higher levels if the color of the tablets and the film are of very different colors
Poor color masking	Reformulate the film coating to a darker color and/or increase the quantity of opacifier
Poor uniformity of solution application	Increase the pan speed
	Reduce the spray rate/increase the coating time
	Increase the spray pattern width
	Increase the number of spray guns
	Check the uniformity of solution delivery through the spray guns (calibration)

7. Tablet attrition/erosion: Attrition is exhibited when some portion of the product substrate exhibits a high level of friability. This typically occurs at the tablet edge or face. Often signs of attrition will be minimal or nonexistent in smaller diameter coating pans. However, when the coating process is scaled-up to a production-sized pan, this problem can become more severe due to the increased batch weight and bed friction[11].

Possible cause	*Remedies*
Insufficient tablet friability	Reformulate the tablet core to a friability of no more than 0.1% friability
Poor film-coating formulation	Reformulate to provide coating with greater mechanical strength
Excessive pan speed	Reduce the pan speed to the minimum required to achieve a smooth and continuous bed movement
Insufficient spray rate	Increase the spray rate to provide a protective film coating in a shorter time, this may require an adjustment in the inlet air temperature

8. Core erosion: Core erosion is another type of attrition due specifically to overwetting of the tablet core. With this type of defect, excessive overwetting may cause a partial disintegration of the core surface.

Possible cause	*Remedies*
Surface overwetting	Reduce the spray rate
	Increase the inlet and exhaust air temperatures
	Increase the spray pattern width
	Reduce the spray droplet size/increase the atomization air volume
	Reformulate the core with less water sensitive excipients
	Improve tablet bed movement
	Increase the bed to gun distance

9. Peeling: Peeling occurs when large pieces or flakes of the film coating fall off the tablet core.

Possible cause	*Remedies*
Poor adhesion	Reduce the amount of insoluble additives in the coating solution
	Increase the level of film former in the coating solution
	Reformulate to incorporate a film polymer with greater adhesion
	Reformulate the tablet core to increase porosity
Brittle film coating	Increase and/or switch to an alternative plasticizer

10. Loss of logo definition: This defect occurs when the tablet logo is no longer clearly legible. It may be due to one or more of the defects previously covered: core erosion, tablet attrition, or bridging. Loss of definition can also occur when the logo is filled in with spray dried film.

Possible cause	*Remedies*
Core erosion	Reduce the spray rate
	Increase the inlet and exhaust air temperatures
	Increase the spray pattern width
	Reduce the spray droplet size/increase the atomization air volume
	Reformulate the core with less water-sensitive excipients
	Improve tablet bed movement

(*Continued*)

TABLE 34.18 (Continued)

Excessive film-coating level	Reduce the film-coating weight gain
	Improve the uniformity of the coating distribution
Tablet attrition	Reformulate the tablet core to a friability of no more than 0.1%
	Reduce the pan speed to the minimum required to achieve a smooth and continuous bed movement
	Increase the spray rate to provide a protective film coating in a shorter time; this may require an adjustment in the inlet air temperature
Bridging	Reformulate the coating solution to improve the film adhesion
	Reformulate the core formulation to increase porosity
	Redesign logo to incorporate shallower angles
Excessive drying rate	Reduce the inlet and exhaust air temperatures
	Reduce the gun-to-bed distance
	Reduce the nozzle air pressure or atomization air volume

11. Core stability issues**:** These problems may manifest as discoloration of the core or degradation of the core active.

Possible cause	*Remedies*
Moisture sensitivity	Explore the remedies for overwetting defects
	Reformulate the core with less water-sensitive excipients
Heat sensitivity	Decrease the inlet and/or product temperature
Tablet component incompatibility	Investigate the compatibility of the tablet core with the coating formulation
	Determine the compatibility of the core ingredients

12. Tablet marking: Usually evident as the presence of black marking on the face of the tablets.

Possible cause	*Remedies*
Abrasion of oxide from the coating pan surface	Change to a tablet shape that tumbles rather than slides (ie, round rather than oval/capsule)
	Reduce the level of titanium oxide in the coating
	Add baffles to reduce the sliding of tablets
	Coat the pan with coating solution to act as a barrier
	Increase the spray rate or reduce the pan speed

34.7 APPLICATION OF SYSTEMATIC AND STATISTICAL TOOLS FOR TROUBLE SHOOTING AND PROCESS OPTIMIZATION

As we move firmly into the 21st century, the QbD, physician quality reporting initiative (PQRI), and risk management are introduced to allow the manufacturer and the regulator to better design, monitor, and control the quality of the product in United States, Europe, Japan, and other countries.

These are some key initiatives:

Concept of design space
QbD
Risk assessment
DOE
Knowledge space/control space
ICH Q8 (development)
ICH Q9 (risk management)
ICH Q10 (quality system)
KPP/CPP
Key quality attribute/critical quality attribute

To meet these challenges, many systematic and statistical tools are being used in a sequential manner to

- Identify root cause(s);
- Understand the KPP/CPP and knowledge space/control space;
- Optimize the process; and
- Establish the control plan, which should be one for continuous improvement.

Many pharmaceutical companies employ a variety of systematic and statistical tools and define the procedures of their approach. These tools include

six-sigma/DMAIC, DOE, Kepner Tregoe analysis (KT), Brightest, Pareto Analysis, Gauge RR, Lean sigma, risk assessment, and other tools.

Case Study: Using Systematic and Statistical Tools by Right-First-Time Approach to Address Film Coating Issues.

PFIZER METHOD 1 RIGHT-FIRST-TIME INVESTIGATION

John Z. Smith (GMS/MPS)
Michael Chen, (NPG/PPD)

DEFINE : Poor Color Uniformity and Uneven Surface of Film Coating for Product XXXXXX.

The film coating of Product XXXXXX shows a poor colour uniformity and rough surface when transfer from Site A to Site B.

MEASURE

Observation:

1) Site A:
Material: a commercial blue film coating material
Coating Pan: fully perforated
Pan Load: 45 kg
Inlet Air Temperature: 60-65°C
Bed Temperature: 40-42°C
Exhaust Temperature: 45°C
Pan Speed: 7 rpm
Number of Gun: 2
Gun to Bed Distance: 25 cm
Spray Rate: 120 g/min total

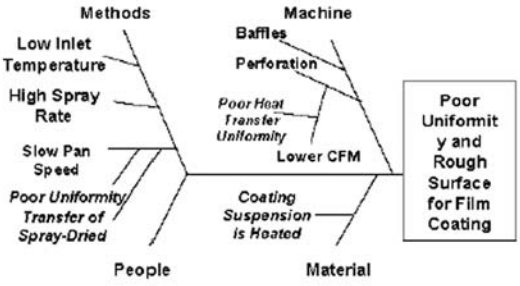
2) Site B
Material: same as Site A
Coating Pan: solid pan
Pan Load: 95 kg
Inlet Air Temperature: 45-50°C
Bed Temperature: 34-36°C
Exhaust Temperature: 40°C
Pan Speed: 10-12 rpm
Number of Gun: 3
Gun to Bed Distance: 30 cm
Spray Rate: 150-250 g/min total

The resulting tablets did not meet the Appearance Quality Limits. About 20% of coated tablets show significant lighter shade. The rough surface appearance show some solids deposits about 500-2000 µ.

ANALYSE

Cause & Effect Analysis (Fish-Bone Analysis)

Cause and Effect Analysis



.

Based on the shape and size of surface defects, some coating materials are detached and transferred onto other tablets. This could due to the imbalance of heat exchange, pan speed, and spray rate,

Critical Process Parameters
1 Bed Temperature
2 Pan Speed
3 Spray Rate
4 Atomization
5 Gun to Bed Distance

IMPROVE

DOE Coating Study
.3X3 Full factorial design
Factor 1: Bed Temperature
Factor 2: Pan Speed
Factor 3: Spray Rate/Atomization

The knowledge space is identified and better understood. The optimized parameters are identified and successful full-scale batches are manufactured.

The final parameters for Site B are listed below:

Material: same as Site A
Coating Pan: solid pan
Pan Load: 95 kg
Inlet Air Temperature: 50-55°C
Bed Temperature: 40-42°C
Exhaust Temperature: 42-45°C
Pan Speed: 12 rpm
Number of Gun: 3
Gun to Bed Distance: 25 cm
Spray Rate: 180-250 g/min total

CONTROL

The control plan is established for future manufacture.

References

1. Porter S, Bruno CH, Jackson GJ. *Pan coating of tablets and granules. Pharmaceutical dosage forms: tablets*, vol. 3. New York, NY: Marcel Dekker; 1982.
2. Liu L. The coating efficiency of a film-coating process. In: *Invited speaker, Pharmaceutical excipients conference*, June 25–27, Philadelphia (PA).
3. Liu L, McCall T, Tendolkar A, Shlyankevich A, Giambalvo D, Baumgamer C, et al. The scale-up and process optimization of a ethylcellulose–based functional coating and its impact on in-vitro/in-vivo release of a novel controlled release system. In: *The 27th international symposium on controlled release of bioactive materials*. Paris (France); July 7–13, 2000.
4. Liu L, Giambalvo D, McCall T, Labudzinski S, Baichwal A. The effect of process variables on coating efficiency of an HPMC-based film coating using a perforated coating pan. In: *AAPS 1998 annual meeting*. San Francisco (CA); 1998.
5. Freers S, Jensen B, Shipley C, Foltz D. *Clear coated tablets with high-solids aqueous starch coating and minimal tablet weight gain*. Vector Corporation.
6. Liu L, Smith TJ, Sackett G, Poire E, Sheskey P. Comparison of film coating process using fully perforated and partially perforated coating pans. In: *2004 AAPS annual meeting* 2004.
7. Bauer K, Lehman K, Osterwald H, Rothgang G. *Coated pharmaceutical dosage form fundamentals, manufacturing techniques, biopharmaceutical aspects, test methods and raw materials*. CRC Press Boca Raton, FL; 1998.
8. Wade A, Weller PJ. *Handbook of pharmaceutical excipients*. 2nd ed. Washington, DC: A.Ph.A; 1994.
9. McGinity JW. *Aqueous polymeric coatings for pharmaceutical dosage form*. 2nd edition. New York, NY: Marcel Dekker, Inc.; 1997.
10. Rowe RC, Force SF. The effect of Process Conditions on the Incidence of Bridging of the Intagliations and Edge Splitting and Peeling on Film Coated tablets. *Acta Pharm Tech* 1982;**28**:207–10.
11. Russo EJ. Typical scale-up problem and experiences. *Pharm Technol* 1984;**8**(1):46–56.

CHAPTER

35

Development, Optimization, and Scale-Up of Process Parameters: Wurster Coating

D. Jones[1,†] *and E. Godek*[2]

[1]Glatt Air Techniques Inc., Ramsey, NJ, United States [2]OWI-Consulting, Inc., Coatesville, PA, United States

35.1 INTRODUCTION

In 1959, Dr Dale Wurster, then at the University of Wisconsin, introduced an air suspension coating technique now known as the Wurster system. The Wurster process enjoys widespread use in the pharmaceutical industry for layering and film-coating of particles and pellets, as well as the emerging controlled-release tablet dosage forms. Product containers typically range in size from 3.5″ (100–500-g batch sizes) to 55″ (up to approximately 1100 kg). The Wurster process is used commercially for coating particles from less than 100 μm to tablets, and for layering to produce core materials.

35.2 BASIC DESIGN

The basic design components of a commercially available Wurster system are shown in Figs. 35.1 and 35.2. The coating chamber is typically slightly conical, and houses a cylindrical partition (open on both ends) that is approximately half the diameter of the bottom of the coating chamber (in up to 18″ Wurster coaters). At the base of the chamber is an orifice plate that is divided into two regions. The open area of the plate under the partition is very permeable. This permits a high volume and velocity of air to pneumatically transport the substrate vertically through the partition. As they accelerate upward, particles pass a spray nozzle that is mounted in the center of this up-bed orifice plate. The nozzle is referred to as a two-fluid or binary type, where liquid is delivered to the nozzle port at low pressure and is atomized by air at a preselected pressure and volume. The spray pattern is generally a solid cone of droplets, with a spray angle ranging from approximately 30–50 degree (Fig. 35.3). The so-called "coating zone," which is formed is a narrow ellipse, and varies in volume depending on the size of the substrate being sprayed and the pattern density in the partition.

The diagram in Fig. 35.4 illustrates the regions of flow in the Wurster process. The region outside of the partition is referred to as the down bed. The configuration of the orifice plate in this area depends on the size and density of the material to be processed. The purpose of the airflow in the down bed region is to keep the substrate in near-weightless suspension, irrespective of its distance to either the wall of the product container or the partition. The goal is to have it travel rapidly downward (to minimize cycling time), and then be drawn horizontally toward, and ultimately into the gap at the base of the partition. In general, larger substrates, like tablets, require significantly more air to produce this condition than smaller substrates, like pellets or fine particles. The orifice plate must be selected accordingly (no single down-bed plate can achieve good fluidization properties for all substrates). As mentioned previously, the material flow inside of the partition is controlled by the up-bed or partition plate. In general, this plate is considerably more permeable than the down-bed plate. In pilot- and production-scale Wurster systems, the up-bed plates are removable so that one or the other may be changed to fine-tune the behavior in these individual regions (see Fig. 35.5).

A second key process variable in Wurster coating is the height that the partition sits above the orifice plate, which controls the rate of substrate flow horizontally

†Deceased.

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00035-2

FIGURE 35.1 Glatt model GPCG-60 fluid-bed processor fitted with an 18″ HS Wurster insert. *Source: Photograph courtesy from Glatt Air Techniques, Inc.*

FIGURE 35.2 18″ HS Wurster coater—inside view showing spray nozzle, HS nozzle surround, and partition. *Source: Photograph courtesy from Glatt Air Techniques, Inc.*

FIGURE 35.3 Spray nozzle and spray pattern (using water).

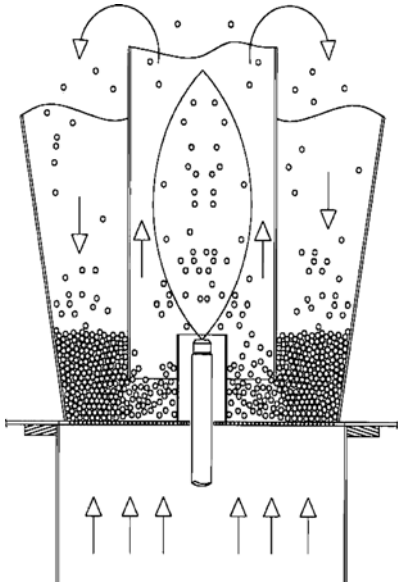


FIGURE 35.4 Diagram of the HS Wurster process.

into the coating zone. Typically, the smaller the particles to be coated, the smaller the gap will be. When the Wurster coating chamber is assembled properly, the resulting flow pattern should be relatively smooth and rapid in the down bed, and very dense and homogeneous in the up-bed (or partition) region.

The substrate exits the partition at a high rate of speed, requiring a region to decelerate. Above the product container is the expansion area (or freeboard, in traditional fluidized bed terms), which is typically conical to allow for decreasing air and particle velocity. Wurster machines designed for pellets and small particles employ elongated expansion chambers to essentially enhance deceleration in the air space rather than by high velocity impact against machine components in the filter housing. By contrast, tablets do not need much expansion height, and in fact, attrition may be a severe problem if they travel too high into the expansion area. The orifice plate and partition height should be optimized so that the tablets travel upward only a very short distance out of the partition before beginning their descent. As a result, a mesh bonnet can be used in the expansion chamber, just above the product container, to keep the tablets from colliding with the expansion chamber. The coarse mesh is intended to allow the fines from the cores or some spray-dried

FIGURE 35.5 Up- and down-bed orifice plates for an 18″ HS Wurster coater. *Source: Photograph courtesy from Glatt Air Techniques, Inc.*

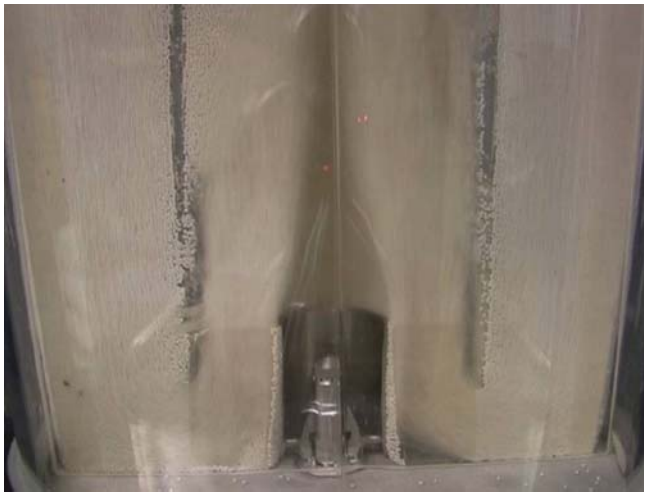


FIGURE 35.6 Spray nozzle and pattern shown with its nozzle surround (patented). *Source: Courtesy from Glatt Air Techniques, Inc.*

coating to exit the process area, avoiding incorporation in the layers of film. A conventional Wurster machine is typically not used extensively for tablet coating because of the comparatively high stress to which the tablets may be exposed. However, it is recommended when the film quality (minimal defects) or active component distribution uniformity are very important, especially for modified-release tablets. The films applied by Wurster systems are high quality due to the concurrent spray and high-drying efficiency of this air suspension process. There have been modifications to the Wurster process specifically for tablet coating. These include spray nozzle surrounds, partition geometry, and airflow adjustability at the interface between the wall of the product container and the down-bed orifice plate (to be described in more detail later in this chapter).

Pellets and small particles are layered or coated extensively via the Wurster process using water, organic solvents, or even by spraying molten materials. All fluidized-bed techniques are known for high rates of heat and mass transfer, and the Wurster process is very effective in this regard. Highly water-soluble materials can be coated using water-based applications without concern for core penetration. Droplets applied to the surface spread to form a continuous film or layer, and then quickly give up their moisture to the warm, dry air. After a thin film has been applied, spray rates can be increased because the soluble core has been isolated. Films applied with volatile organic solvents are also high in quality because the formed droplets impinge on the substrate very quickly, minimizing the potential for spray drying of the film.

35.3 HS WURSTER CONSIDERATIONS

HS Wurster technology, shown in Figs. 35.4 and 35.6, is a product of Glatt Air Techniques, Inc., in Ramsey, NJ, and involves the use of a proprietary device to influence the behavior of the substrate in proximity to the coating zone. Unfortunately, the liquid spray-application rate is not controlled by the drying capacity of the fluidization air, but by the nature of the coating material (tackiness) or by the region immediately surrounding the spray nozzle (the first area of focus is this region). In all Wurster inserts, the high velocity of air rushing through the partition relative to the lower velocity in the down bed creates suction at the partition gap according to Bernoulli's Principle. As a consequence, particles entering the partition through the gap have a horizontal component to their flow. Some travel toward the spray nozzle instead of simply making the transition from horizontal to vertical flow. The atomizing air has a very high velocity (it is likely supersonic), and creates another area where Bernoulli's Principle applies, which draws the substrate to the base of the developing spray pattern. In a standard Wurster machine, the nozzle is elevated above the orifice plate at a position just below the confluence of fluidizing substrate particles. A portion of the developing up-bed stream of product can pass the nozzle tip, either closely or at a further distance. The zone directly above the nozzle tip is where the coating liquid and the atomizing air mix, and droplets begin to form. Particles that are very close to this zone tend to be strongly overwetted, and if they contact other particles, agglomerates are formed. To control what could otherwise be severe agglomeration, a typical response is to reduce the spray rate. This leaves a large amount of the drying capacity unused. Other commonly used agglomeration control techniques include raising the inlet (and product) temperature to increase the drying rate or raising the atomizing air pressure to shrink droplet size—options which are in conflict with producing high-

quality films. Even using these corrective measures, a quantity of agglomeration in traditional Wurster systems is almost inevitable.

The HS modification for the Wurster process was conceived to keep particles away from the spray nozzle until the spray pattern is fully developed. As a result, more of the excess drying capacity can be used, and the application rate increased substantially (more than doubled in many pilot-scale experiments). Agglomeration is also substantially diminished or eliminated because the particles are kept away from the wettest portion of the pattern (Fig. 35.6). An additional benefit is that the high-atomizing air velocities necessary to produce very small droplets for coating of particles smaller than 100 μm may be useable without pulverizing the substrate. The velocity of this air diminishes dramatically with distance from the nozzle, and even a few centimeters are significant. Therefore, keeping the product away from the nozzle tip allows the atomization air velocity to decrease significantly before contacting the substrate, reducing the likelihood of attrition, especially during the early stages of coating.

Coating of substrates smaller than 100 μm has been achieved more frequently using the HS Wurster coater. Success depends on many factors, both process- and product-related. Product considerations, such as flow properties of the substrate (generally poor in this size range, which must be improved), as well as the liquid, which must be amenable to atomizing to droplets well below 10 μm, must be addressed. The tremendous surface area of such fine particles also requires very high coating quantities, and consequently, a low potency of the final-coated product (often less than 50%).

35.4 COATING AND PROCESS CHARACTERISTICS

The coating liquid is sprayed in the direction of motion of the fluidizing particles. In general, the fluidization is orderly, with very rapid, dilute-phase pneumatic transport in the up bed, and relatively smooth and rapid transport in the down-bed region (outside of the partition). Droplet travel distance is minimized because the liquid is sprayed into a well-organized pattern of substrate moving relatively close to the nozzle. In this manner, droplets reach the substrate prior to any appreciable evaporation. By retaining their low viscosity they are able to spread on contact, and the resultant films are excellent, even when using organic solvents as an application medium. The drying efficiency of the fluidized-bed also minimizes the potential for core penetration, and the sample shown in Fig. 35.7 clearly shows a well-defined boundary layer between the substrate and the coating material; in this case a film applied using an aqueous dispersion.

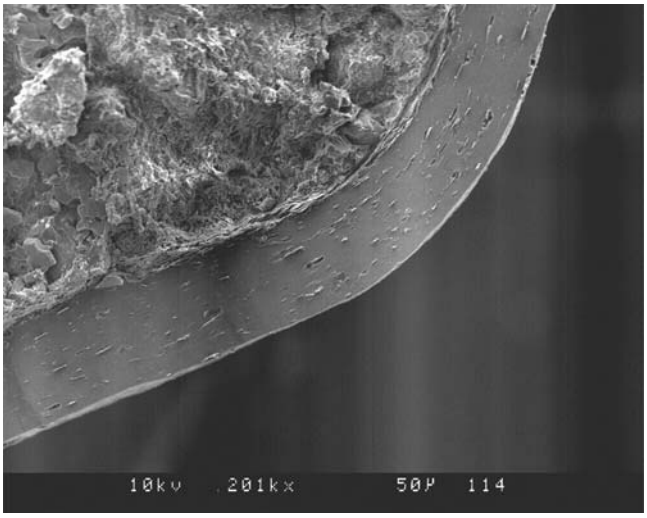


FIGURE 35.7 Cross-section showing boundary layer between coating and core material (200× magnification).

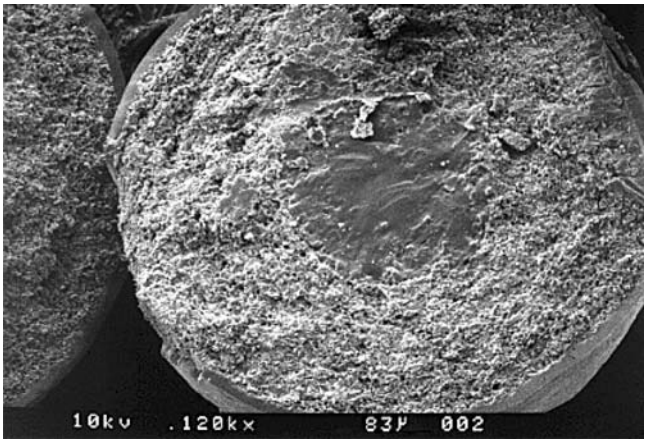


FIGURE 35.8 Cross-section of drug-layered, film-coated pellet (120× magnification).

35.5 PROCESSING EXAMPLES

The Wurster process is used commercially for coating and layering (high solids build up onto a type of core material). Substrates include particles smaller than 100 μm, crystals, granules, pellets, and tablets. As mentioned previously, films may be applied using water-based solutions, aqueous dispersions, or organic solvents.

The first example of a material layered and coated using the Wurster process is shown in cross-section in Fig. 35.8. The starting material is a very small, nonpareil sugar seed, and it is layered to a high potency using a suspension of drug in a dilute-binder solution. Release of the drug is controlled by a film subsequently applied. All three components are clearly identifiable.

The example shown in Fig. 35.9 is a tablet, film-coated (for sustained release) and layered with an additional dose of active material prior to drilling with a laser. Immersed in liquid, drug solute exits the laser hole in this type of dosage form. The Wurster process was selected for this product due to its abilities in film and drug distribution uniformity.

The product in Fig. 35.10 is coated with a moisture barrier. The substrate contains a significant quantity of material smaller than 5 μm. Agglomeration of fine particles does occur, as seen in the scanning electron micrograph (SEM). However, the average particle size of the finished product is still less than 50 μm. The film is applied from a volatile organic solvent solution, and is typically very tacky, hence the agglomeration "crater" where another particle has detached.

A water-soluble core, coated with a soft, heat-sensitive polymer system is shown in Fig. 35.11. The cross-section shows the deformable nature of the film, and these particles may be filled into capsules or compressed into tablets to be delivered. Although the product must be processed at low temperatures, core penetration was not a problem due to the process parameters used and the drying capability of the Wurster process.

Often the economics of coating very small particles dictate whether a product ever reaches the marketplace. When coated in a standard Wurster system, the drug particles shown in Fig. 35.12 proved too costly to produce. However, when the HS components were fitted to the pilot-scale 18″ Wurster machine, the process time was reduced by more than 50%, and the product is now commercially produced. The SEM shows very good morphological properties for the core, in that the particles are nearly spherical, and are not porous. In laboratory trials, this substrate was found to be very robust—attrition of the cores early in the coating process (before they were sufficiently coated to have improved surface strength) was essentially nonexistent.

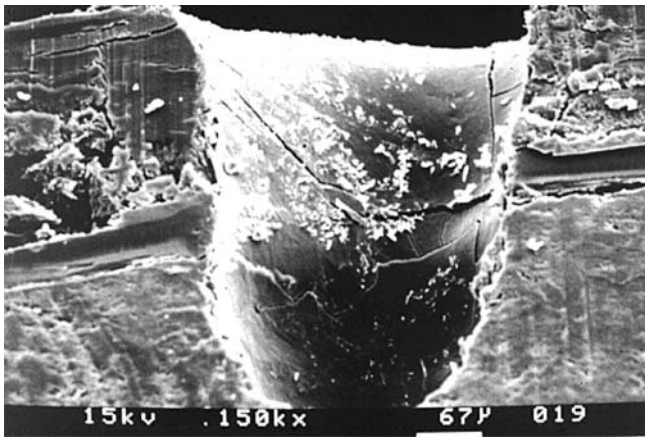


FIGURE 35.9 Cross-section of a tablet coated using the Wurster process (150× magnification).

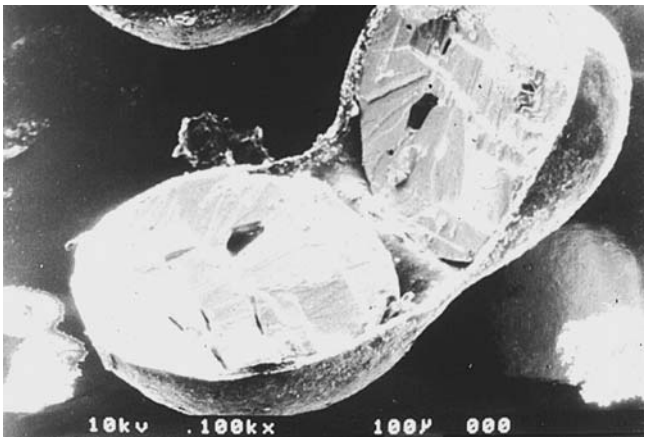


FIGURE 35.11 Water-soluble crystal coated for sustained release (100× magnification).

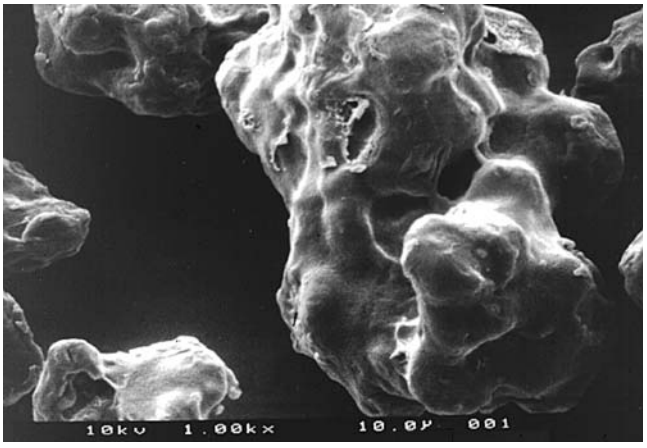


FIGURE 35.10 Very fine particles coated with a moisture barrier using the HS Wurster process (1000× magnification).

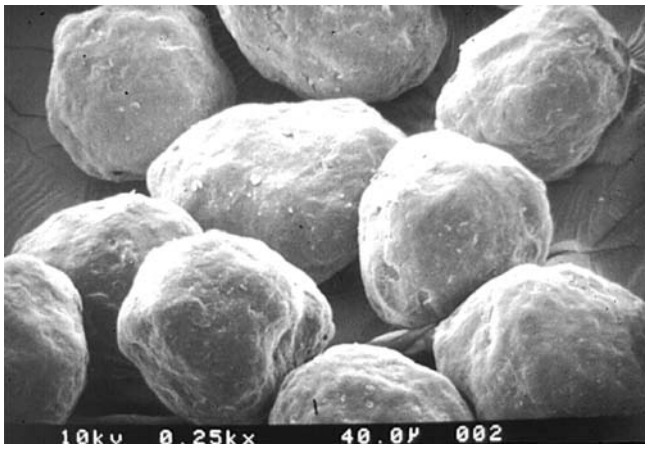


FIGURE 35.12 Drug core particles coated for taste masking, incorporation into oral suspension (250× magnification).

35.6 PROCESS VARIABLES

35.6.1 Batch size

The working capacity for a Wurster insert of a particular size is generally defined as the volume outside of the partition, with the partition at rest on the orifice plate. Loading of product only in the area outside of the partition helps to ensure that the nozzle remains clear at the beginning of the process. This is especially important for fine particle or pellet coating. For tablet coating, up to half of the volume inside the partition can also be used. As the bed becomes fluidized, the dense phase or surface of the down bed will drop, indicating that there is room for the batch to expand as coating is applied. In some cases, it may appear that there is more than enough, and that batch size could easily be increased. The danger is that the level of the bed is dynamic—varying in depth in response to the behavior of the up bed and surface properties of the substrate during spraying. Should the substrate begin to drop down into the partition against the desired flow, the up bed would be seriously disturbed, agglomeration would be likely, and the batch might be lost.

The minimum batch size is approximately 20–25% of working capacity. It is critical that there be sufficient material in the up-bed region to accumulate all or most of the coating material being sprayed or "efficiency" (actual yield vs expected) will suffer. The down bed acts as a material reservoir, "feeding" the base of the partition (entrance to the coating zone), and if its depth is insufficient, the up bed will be too sparse, favoring spray drying. In a more extreme case, coating of the inner wall of the partition may be experienced (this is more common in very small laboratory-scale inserts). Use of such a small volume of the product container is recommended only if the final batch size is significantly larger than the starting volume (eg, layering). Early in the process, the coating efficiency is compromised. This inefficiency diminishes as the batch size increases, improving mass flow through the coating zone. When approximately half of the down bed working capacity is occupied, the efficiency will be at or near its maximum for the formulation being executed. Ultimately, the inefficiency seen at the beginning of the process should be of little or no consequence to the finished batch.

After the layering step, it is strongly recommended that the minimum batch size be increased to approximately 50% of the working volume if a thin film is to be applied. For film-coating, inefficiency will likely impact the resultant dissolution behavior. Some amount of spray drying may result in unexpectedly rapid release of drug, particularly for a sustained-release coating.

35.6.2 Fluidization pattern

Irrespective of the amount of substrate in the Wurster insert, the fluidization pattern is controlled by the orifice plate configuration, the partition height, and the fluidization air volume. As mentioned previously, the goal is to have a rapid and relatively smooth down bed. This is not always possible, depending somewhat on particle size (small particles are difficult, if not impossible, to fluidize without bubbling), but to a greater extent on the tackiness of the product during spraying. Some coating materials are inherently tacky or move sluggishly due to temperature sensitivity. If this is the case, it is more important to keep the entire bed moving and involved by marginally over-fluidizing (turbulence or bubbling in the down bed), rather than attempting to achieve a smooth flow. This can be accomplished by raising the partition somewhat or using a more permeable down-bed plate, to allow extra air to percolate through the descending product. The risk in not doing this is that, as the process progresses, the bed may stall in some region on the orifice plate, and a portion of the batch may not receive all of the desired coating. Typically, this happens in the back portion of the product container, usually out of sight from the operator. The reason for this is that, in most installations, the process air enters the inlet plenum via a duct behind the machine. More air enters the product bowl in the front, due to momentum because the air has mass and needs to make a sharp turn upwards. This can be minimized by increasing the pressure drop across the bottom plate using a retention screen with a lower porosity.

In general, tablets require a large air volume in the down bed to keep it well-aerated. The partition region of the plate is also selected to allow only sufficient air flow to minimize the distance tablets travel out of partition to avoid attrition. By contrast, the orifice plates for pellet coating may differ significantly. The plate is permeable inside of the partition, and the down-bed section is only perforated to the extent that the product flows downward in near-weightless suspension. Achieving this condition typically requires only about a one-fourth of the open area needed for tablet coating. Finally, plates designed for powders are even less permeable in the down-bed region (and boiling may be impossible to control, due to the flow properties of fine materials).

Another critical process variable in the fluidization pattern for Wurster processing is the height of the partition above the orifice plate. The height is also substrate-dependent. For instance, tablets, a comparatively large substrate, would require a substantial partition height, such that a sufficient number of them could pass through the gap per unit time to accept the

TABLE 35.1 Partition Height as a Function of Substrate and Wurster Coating-Insert Size

Substrate	Partition height
Tablets	25–50 mm in small machines
	50–100 mm in 18″ Wursters and larger
Pellets	Approximately 15–25 mm small machines
	Approximately 35–60 mm in 18″ Wursters and larger
Powders	Approximately 25–50 mm in all machines

maximum amount of coating being sprayed into the partition area. Pellets and intermediate particles do not need as high a partition height as tablets, and fine particles and powders need only a small gap to result in an ideal fluidization pattern. Table 35.1 profiles suggested conditions for partition height as a function of substrate and Wurster coating-chamber size.

35.6.3 Atomizing air pressure and volume

Droplet size should be small, relative to the particle size of product to be coated. For example, coating of tablets may need less than 2 bar (30 psi) pressure. Higher pressure and air volume will result in a higher atomization air velocity, increasing kinetic energy at the interface between the spray pattern and slower-moving substrate. The potential for causing attrition of the tablets by accelerating them into machine components is also enhanced. Finally, the high atomizing air velocity may distort the fluidization pattern.

When coating small particles, a somewhat higher atomization air pressure may be necessary, to achieve small droplet size and thereby avoid agglomeration. There is some risk, however, that the high shear associated with pressures in the 3–6 bar (45–90 psi) range, depending on the type and size of the spray nozzle, may cause breakage of fragile core material. The Gustav Schlick Company in Coburg, Germany, is a supplier of spray nozzles in widespread use in Wurster processing. The 970 series nozzle is found in small (3.5″, 4″, 6″, 7″, and 9″) Wurster bottom-spray coaters. With water-like materials, it is useable in a spray-rate range of approximately 0–100 g/minute. Fig. 35.13 shows the influence of atomization air pressure on mean droplet size for water sprayed at 25 g/minute (data by Schlick). Interestingly, increasing the pressure beyond 2 bar does little to decrease the droplet size. This is, in part, due to the fact that 25 g/minute is well within the nozzle's ability to atomize. What should be noted is that, if a process is being run in which agglomeration is a minor problem and increasing the atomization air pressure seems to improve the situation, it is likely a consequence of the increased air velocity and kinetic energy, not a smaller droplet size.

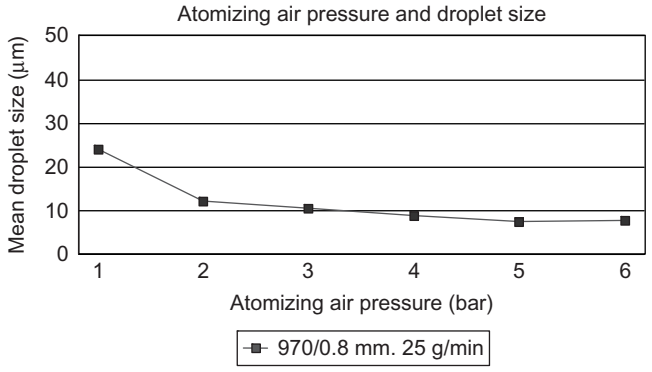
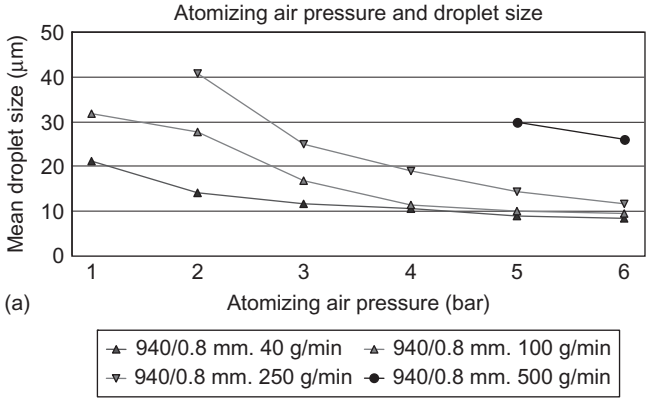


FIGURE 35.13 The relationship between atomizing air pressure and droplet size for a Schlick 970 series nozzle spraying water.



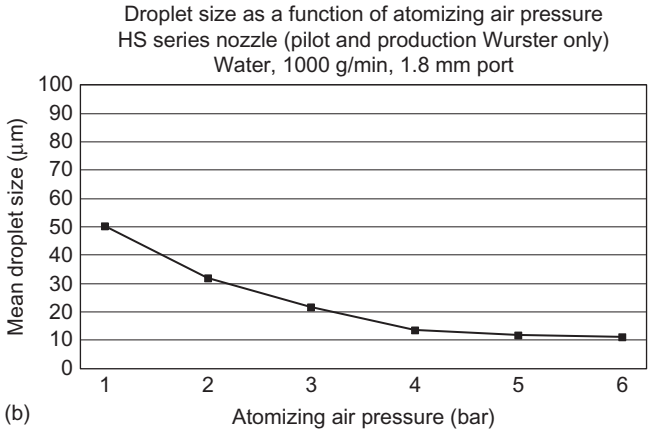


FIGURE 35.14 (a) The relationship between atomizing air pressure and droplet size for a Schlick 940 series nozzle at various spray rates using water. (b) The relationship between atomizing air pressure and droplet size for a Schlick HS Wurster nozzle at 1000 g/min spraying water.

Looking at the performance envelope of the Schlick 940 series nozzle (Fig. 35.14), which is used in older style 12″, 18″, 24″, 32″, 46″, and 55″ Wurster coaters, it can be seen that droplet size increases with faster spray

rates (at a constant atomization air pressure). In cases where the spray rate is 250 g/minute or less, it is possible to increase atomization air pressure/volume/velocity to achieve droplets smaller than 20 μm. However, the data shown for the 500 g/minute rate demonstrates that even at the highest practical atomization air pressure (6 bar), it is not possible to produce 20 μm droplets (spraying water). This is an important consideration in larger capacity equipment, where there may be significant drying capacity, and the rate-limiting factor is the inability of the nozzle to atomize liquid (to a satisfactory droplet size) at the rate at which the process air may remove the resultant water vapor. The only possibility for taking advantage of the increased drying capacity is to enlarge the nozzle (use more compressed air at the same pressure).

The 940 series nozzle is unable to produce droplets smaller than 20 μm at 500 g/minute using water, even at very high atomization air pressures. A process that has excessive drying capacity, but is limited by droplet size (eg, fine particle coating), will result in unnecessarily hindered productivity. Upgrading to the HS nozzle, which uses substantially more compressed air at the same atomization air pressures (approximately three times the volume of the 940 series nozzle), will result in a dramatic improvement in drying capacity utilization. The graph in Fig. 35.14b depicts a droplet profile that is similar to the 940 series nozzle spraying at 250 g/minute, with the exception that the HS data is for 1000 g/minute, a spray rate four times that of the 940 series nozzle. This permits the HS nozzle to be operated at low atomizing air volumes and pressures, limiting the potential for attrition due to high kinetic energy.

35.6.4 Nozzle port size

Droplet size is rarely influenced by nozzle port size. At a fixed volumetric spray rate, the velocity of the liquid into the atomization air is controlled by the nozzle port size. The lower the velocity, the more complete the atomization of the liquid will be, and the smaller the mean droplet size; Fig. 35.15 illustrates this data for a 940 series nozzle for two different spray rates. For 100 g/minute, the mean droplet size is nearly the same for all data points, with the exception of the values at 2 bar. However, this appears to be an anomaly because the mean droplet size is the same for both port sizes at the lower pressure of 1 bar. At the higher spray rate of 250 g/minute, the mean droplet sizes are the same at 3 bar and higher. Only at 2 bar is there an appreciable difference, owing to the lower liquid velocity, and longer dwell time for the liquid in the atomizing air stream. In addition, factors such as liquid viscosity and surface tension influence this behavior; and it is preferable that they both be low to avoid the need for high atomization air pressure, regardless of the size of the substrate. If viscosity is an issue (high), nozzle port size should be selected to accommodate it (to minimize back pressure in the spray pump).

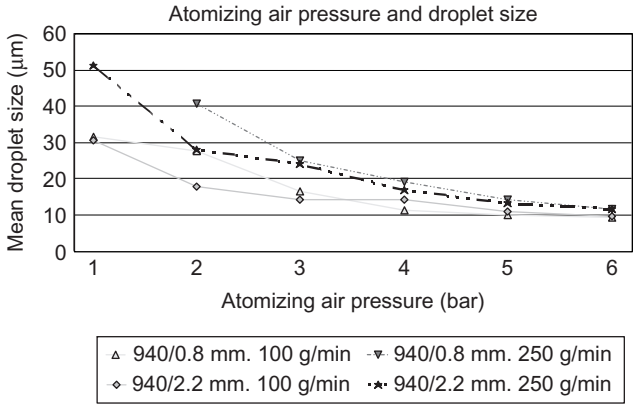


FIGURE 35.15 The relationship between atomizing air pressure and droplet size for a Schlick 940 series nozzle at 100 and 250 g/min (water) using 0.8 and 2.2 mm nozzle port sizes.

35.6.5 Evaporation rate

Evaporation rate or drying capacity is controlled by fluidization air volume, temperature, and its absolute humidity (dew point). The temperature of the fluidization air is generally adjusted to maintain a constant product temperature. The product temperature, in turn, is influenced by spray rate and humidity of the incoming air. First, to take advantage of excess drying capacity, inlet temperature must be raised to keep the product temperature from dropping, due to the increase in evaporative cooling as spray rate is increased within a given batch. Secondly, if the dew point of the incoming air is not controlled, the drying conditions to which the product is exposed will not be reproducible. The dew point changes seasonally, especially in northern climates, and may range from below 0°C during winter to more than 20°C in summer. These normal variations in weather conditions may result in some notable influences in processing.

In cold weather, static electricity during fluidization may be a problem, and in summer, high humidity conditions may result in substantially slower spray rates, due to erosion of drying capacity, especially for aqueous coatings. This phenomenon is, of course, more pronounced when product temperatures of 30°C and less are desired (thermally sensitive substrates or coating materials). Allowing the process air dew point to vary will result in a variation in the residual moisture in the applied film. This can lead to a change in the glass transition temperature for some coatings.

For latex materials, residual moisture may result in "aging" effects—changes in dissolution with time (product instability). For most products, a dew-point range of approximately 10−20°C throughout the year is recommended, to minimize this so-called "weather effect." If the product is very sensitive to moisture in the process air or even thermally sensitive, a desiccant dryer is suggested, which will produce dew points well below 0°C. In any case, process air dewpoint control is recommended in the system's air handling unit.

35.6.6 Product temperature

All air suspension processes are noted for their high heat and mass transfer capability. A wide range of product temperatures may be used. Values below 30°C may be used for heat-sensitive materials. Values exceeding 50°C may actually be used with little or no spray drying of the applied coating. High product temperatures are especially attractive for water-based spray liquids. The rate-limiting step for most processes is related to the physical properties of the liquid being applied. The use of psychrometry is effective in determining the threshold at which moisture applied to the surface of the substrate begins to remain in the applied layer, rather than to leave with the process air. This inflection point is where agglomeration begins to take place. The surface tackiness promotes the sticking together of substrate particles. This so-called exit air relative humidity threshold is related to the product temperature. For example, for an exit humidity threshold of 60%, the water removal rate for a product temperature of 40°C is approximately 100% higher than for a product temperature of 30°C. Water content in air is an exponential function—for high product temperatures, the water removal rate is dramatically increased without fear of agglomeration.

35.7 CASE STUDIES FOR LAYERING AND FINE PARTICLE COATING

With an understanding of the range of application of the Wurster process, plus the process and product variables that influence its performance, it is interesting to look at some case studies. The first involves the use of an 18″ HS Wurster coater for solution layering (refer to Figs. 35.16 and 35.17 for in-process data). A water-based drug solution was made, using a high solids concentration. The goal of the trial was to find the maximum possible spray and production rates, avoiding agglomeration as much as possible. The starting material is very fine, initially requiring a reduced spray rate (400 g/minute). However, as the particles grow larger, the spray rate can be elevated as the surface area of individual particles increase. The finished particle size is less than 200 μm, and the potential for productivity using the Wurster process is clearly evident here. The majority of the spray liquid is applied at 900 g/minute. Heat and mass-transfer rates are very high. The inlet air temperature is more than 90°C, the product temperature is approximately 35°C, and the

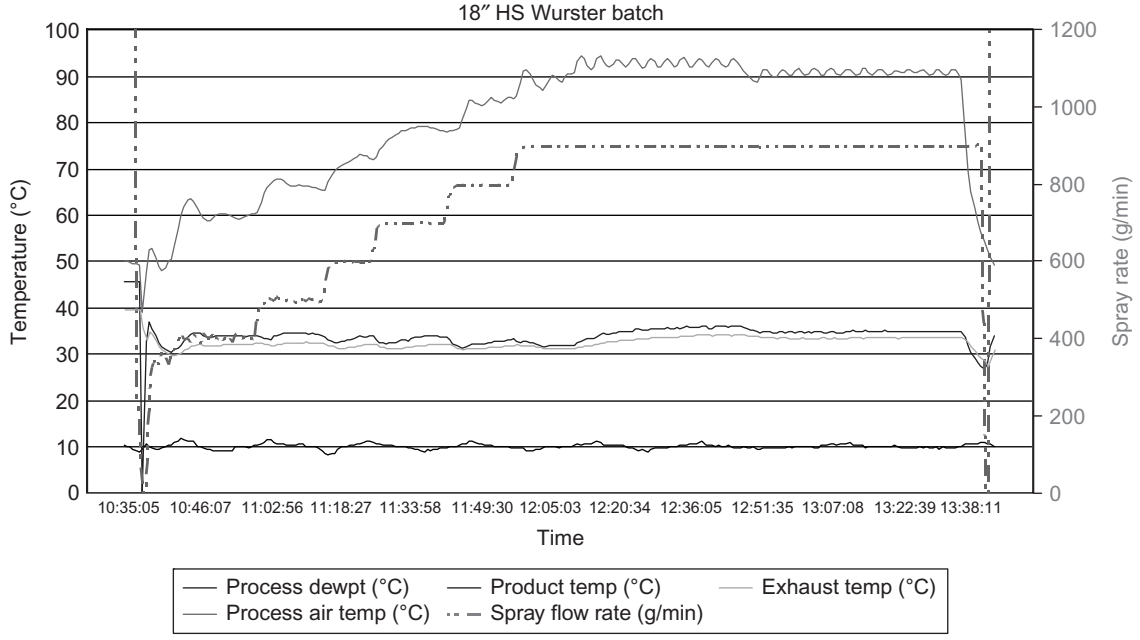


FIGURE 35.16 In-process data (temperatures and spray rate) for an 18″ HS Wurster solution layering process.

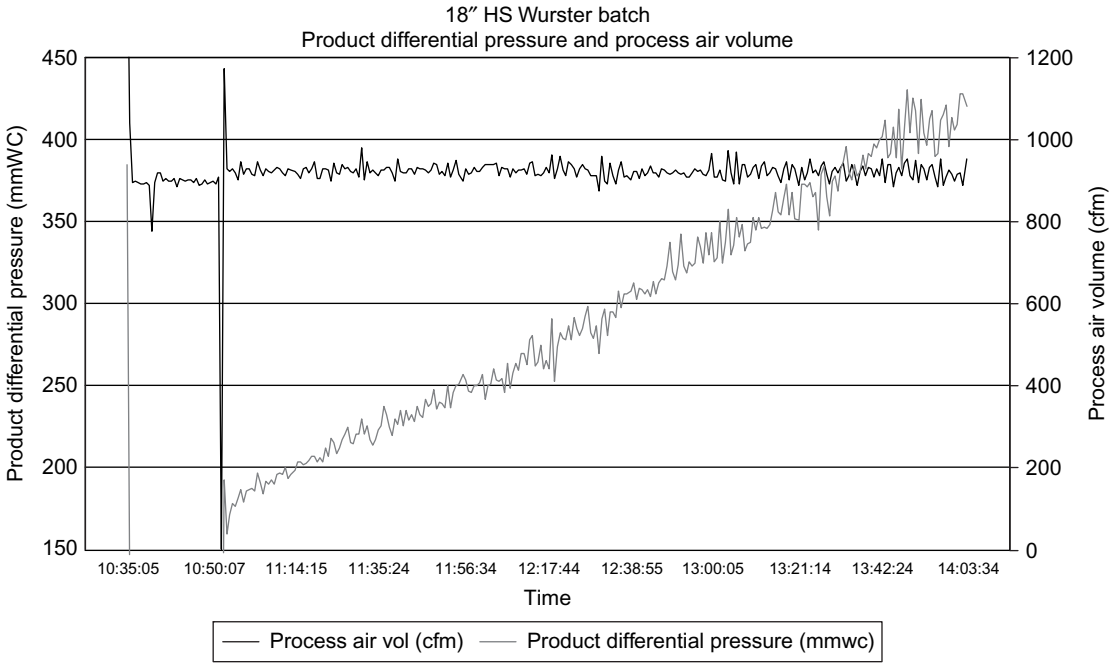


FIGURE 35.17 In-process data (process air volume and product differential pressure) for solution layering in an 18″ HS Wurster coater.

process air volume is 900 cfm. Using psychrometry software, the exit air humidity is found to be at 85% relative humidity, indicating that the vast majority of available drying capacity is being used. An interesting tracing is that of the product differential pressure, shown in Fig. 35.17, as the starting batch size is approximately one-quarter of the finished product weight. As bed depth increases in the down bed, the particles react differently to the up bed process air flow. There is a near-absence of oscillation in the early readings because the down bed is shallow, and there is minimal coalescence of air bubbles as they enter and pass through the bed. However, as batch size increases, the down bed deepens. Small air bubbles, entering through the orifice plate and product retention screen, coalesce into larger bubbles. Back-flow or turbulence in the down bed becomes evident. The escalating peak-to-trough values indicate bubbling air flow in the down bed. This is of no consequence to the process, provided that the down-bed cycling time remains rapid, so that the mass flow in the "coating zone" continues to be high in density and velocity. A consequence of adverse fluidization properties would be agglomeration, however, in this particular pilot-scale trial, this was essentially nonexistent.

The second case study involved coating fine granular acetaminophen particles for sustained release. The uncoated crystals, shown in Figs. 35.18 and 35.19, illustrate a good substrate for coating. They are uniform in size distribution, and do not contain an appreciable amount of fines. What is not seen by SEM, however, is

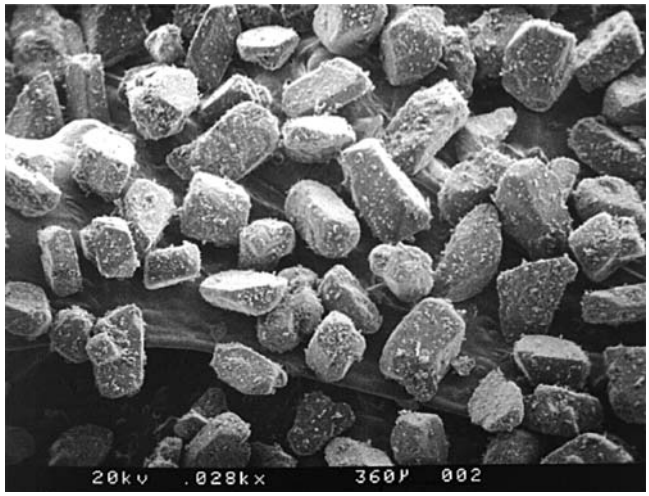


FIGURE 35.18 Uncoated acetaminophen (APAP) crystals (28× magnification).

their fragility. As the particle size data shown in Tables 35.2 and 35.3 reveals, the material is very prone to fracture. Within a few minutes of product warm-up (prior to spraying), the content of material smaller than the 80 mesh screen (180 μm) escalates from nothing to more than 3%, and continues to climb early in the spraying process until sufficient coating has been applied to improve the strength of the primary core material. It is important that the fines that exist in the initial product, in addition to those generated during the early stages of coating, are treated in one of two

ways. First, using a comparatively high spray rate initially, they may be incorporated in the lower layers of film, such that the majority of the remaining film governs the coated-product release properties. Second, an outlet filter media can be selected such that fines are permitted to exit the processor (to be collected in a remote dust collector). If attrition is serious or if fines are continually incorporated in the developing film, as was the case using the conventional 18″ Wurster coater, the release properties of the film will be affected, and in all likelihood, challenging to reproduce. In Figs. 35.20 and 35.21, the acetaminophen crystals have been coated to a 28% coating level (final potency of 78.1%) with AquaCoat (FMC), a pseudolatex of ethylcellulose. The standard version 18″ Wurster equipment was used in a Glatt model GPCG-60 fluid-bed processor, using the HS nozzle, but no HS nozzle surround, and a (then) contemporary version of an up-bed plate referred to as "high mass flow." This plate may be described as having a gradient of holes (and permeability) from the outer perimeter to the spray nozzle. The highest permeability is in close proximity to the nozzle. The intent of this design was to enhance the pressure gradient at the partition gap and increase the flow of material into the coating zone. What became apparent during these trials was that this resulted in a high concentration of substrate close to the tip of the nozzle and the developing spray

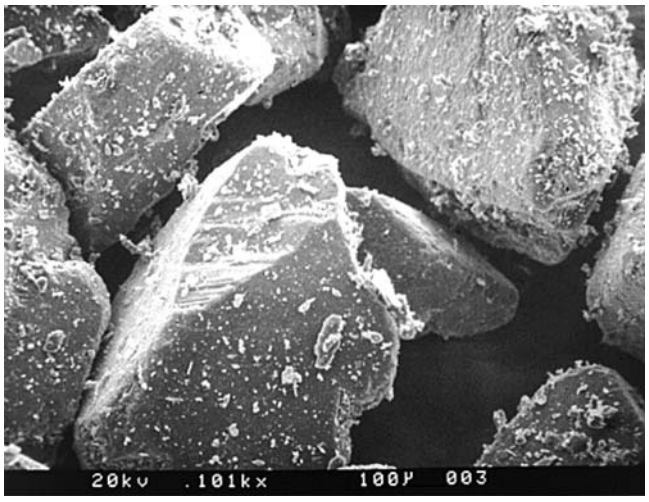


FIGURE 35.19 Uncoated acetaminophen crystals (100× magnification).

TABLE 35.2 Sieve Analysis (During Spraying Without Nozzle Surround)

Time	+30	+40	+60	+80	+100	Pan	Comments
0 min	0.0038 g	2.691 g	33.247 g	6.154 g	0.029 g	0.018 g	
	0.01%	6.39%	78.89%	14.6%	0.07%	0.04%	Starting material
3 min	0.0011 g	0.0658 g	23.441 g	4.519 g	0.327 g	0.461 g	
	0.04%	2.24%	79.69%	15.36%	1.11%	1.57%	Start spray
24 min	0.013 g	0.151 g	15.670 g	6.553 g	1.121 g	2.018 g	
	0.05%	0.59%	61.39%	25.67%	4.39%	7.91%	Pressure from 3.5 to 2.7, increased spray rate from 200 to 270 g/min
37 min	0.031 g	0.561 g	15.873 g	6.729 g	1.661 g	3.828 g	
	0.11%	1.96%	55.34%	23.46%	5.79%	13.35%	Pressure to 2.2 bar, increased spray rate from 270 to 325 g/min
58 min	0.062 g	3.322 g	17.098 g	7.166 g	1.459 g	3.852 g	
	0.19%	10.1%	51.9%	21.7%	4.4%	11.7%	325 g/min
75 min	0.074 g	3.413 g	15.903 g	6.172 g	1.366 g	3.880 g	
	0.24%	11.1%	51.6%	20.0%	4.4%	12.6%	325 g/min
87 min	0.128 g	5.604 g	14.852 g	5.331 g	1.304 g	3.289 g	
	0.42%	18.4%	48.7%	17.5%	4.3%	10.8%	325 g/min
128 min	0.563 g	9.305 g	13.806 g	5.137 g	1.305 g	2.267 g	
	1.77%	29.2%	43.3%	16.1%	4.3%	7.1%	325 g/min
178 min	2.120 g	10.63 g	11.226 g	4.124 g	0.988 g	1.067 g	
	7.0%	35.2%	37.2%	13.7%	3.3%	3.5%	28% coating

Machine configuration: 18″ Wurster coater, HS nozzle, High-mass flow-up bed plate, no nozzle surround, initial atomizing air pressure 3.5 bar.

TABLE 35.3 Sieve Analysis (During Spraying With Nozzle Surround)

Time	+30	+40	+60	+80	+100	Pan	Comments
0 min	0.0065 g	1.358 g	29.037 g	5.266 g	0.24 g	0.008 g	
	0.02%	3.8%	81.%	14.8%	0.07%	0.2%	Starting material
4 min	0.0052 g	0.094 g	26.846 g	3.747 g	0.246 g	0.787 g	
	0.02%	0.30%	84.6%	11.8%	0.8%	2.5%	Begin spray
33 min	0.0061 g	0.223 g	23.958 g	4.701 g	0.749 g	1.451 g	
	0.02%	0.72%	77.1%	15.1%	2.4%	4.67%	11.5 kg applied
68 min	0.020 g	1.382 g	22.495 g	5.619 g	0.826 g	1.614 g	
	0.06%	4.3%	70.4%	17.6%	2.6%	5.1%	26 kg applied
99 min	0.0359 g	3.500 g	20.138 g	4.651 g	0.749 g	1.018 g	
	0.12%	11.6%	66.9%	15.5%	2.5%	3.4%	39 kg applied
130 min	0.0902 g	7.147 g	24.915 g	5.297 g	0.703 g	0.759 g	
	0.23%	18.4%	64.0%	13.6%	1.8%	2.0%	28% coating

Machine configuration: 18″ HS Wurster coater (with nozzle surround), "B" down bed, "G" up-bed plate, atomizing air pressure 3.5 bar throughout.

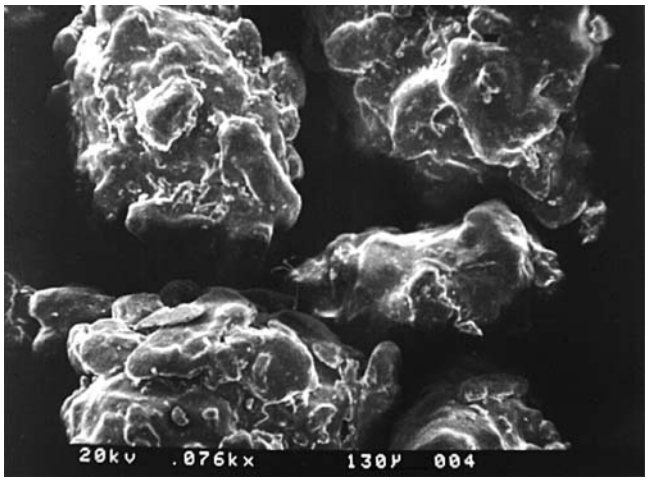


FIGURE 35.20 Acetaminophen—28% coating using standard 18″ Wurster coater (76× magnification).

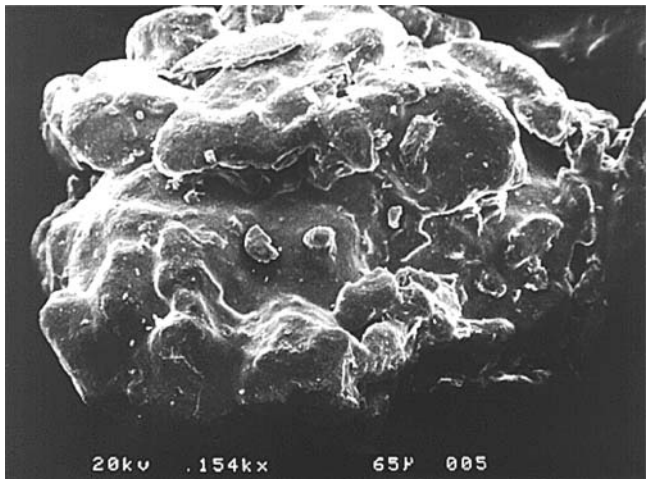


FIGURE 35.21 Acetaminophen—28% coating using standard 18″ Wurster coater (154× magnification).

pattern. Recall that this region is high in liquid content and atomizing air velocity, which can lead to attrition of fragile substrates.

The HS nozzle, for high-capacity spray rates, was used so that the influence of the perforated plates and HS nozzle surround could be seen independently of a key variable, such as nozzle type. In the SEMs, at both the low and higher magnifications, the fines generated early in the spray process due to the high velocity atomization air are readily seen embedded in the film. They are present to the extent that the final particle size is substantially larger than the starting material (see Table 35.2). The resultant porous nature resulted in a faster drug-release rate (Fig. 35.22). This occurs in spite of the fact that the total surface area per unit mass is expected to be lower, due to the increased particle size. This type of problem should be resolved or batch-to-batch reproducibility (dissolution) will be difficult to achieve. In this example, the atomizing air pressure had to be reduced to as low as 2.2 bar to slow the generation of fines, which ultimately became embedded in the film.

In Figs. 35.23 and 35.24, the partition plate was replaced with a more uniformly permeable type "G," and the HS nozzle surround was installed. The surround acts to keep the particles away from the highest velocity and densest droplet region of the spray pattern, as described previously. The result is that the atomization air pressure could be maintained at 3.5 bar throughout the spraying process, with only minimal

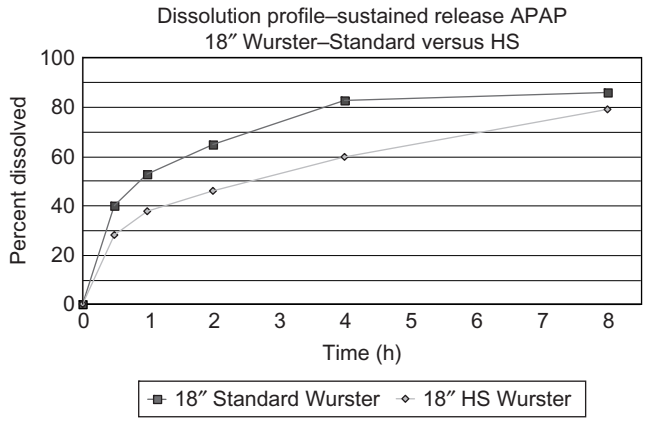


FIGURE 35.22 Dissolution profile for fine granular acetaminophen batches coated in 18″ standard and HS Wursters with 28% w/w coating of AquaCoat ECD (FMC BioPolymer).

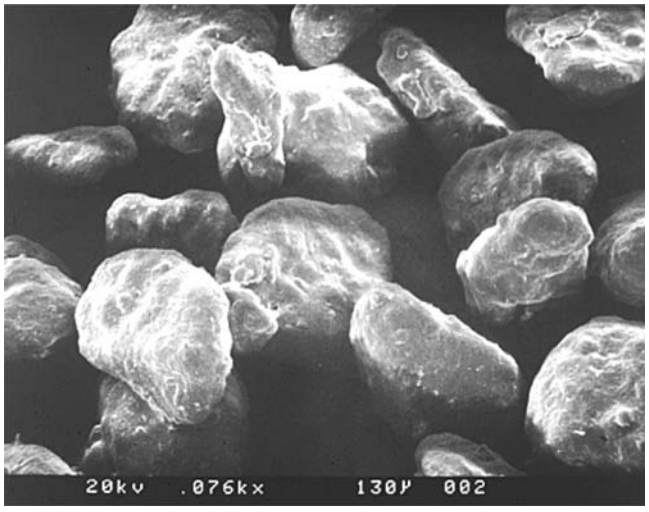


FIGURE 35.23 Acetaminophen—28% coating using 18″ HS Wurster coater (76× magnification).

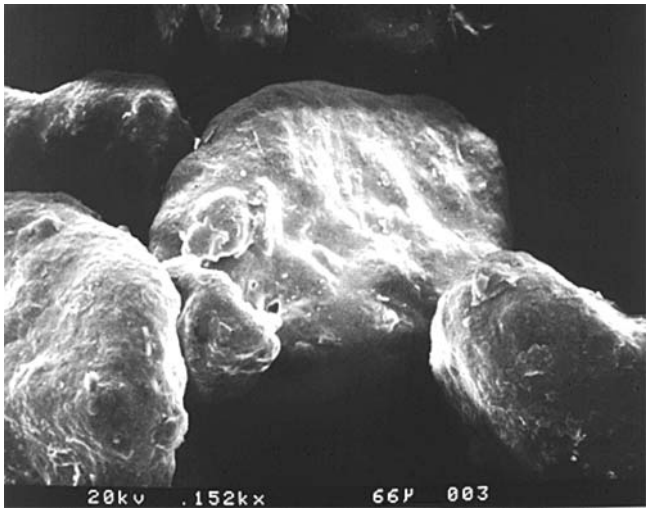


FIGURE 35.24 Acetaminophen—28% coating using 18″ HS Wurster coater (152× magnification).

generation of fines, as evident in the SEMs. The average particle size after coating was smaller than when using the standard Wurster configuration because fewer fines were generated and embedded. This resulted in a less porous film, and slower drug release, as seen in the accompanying dissolution profile shown in Fig. 35.22. An additional benefit was a reduction in process time, due to the ability to apply the coating at a higher spray rate (approximately one-third faster).

For the experiment involving the standard Wurster system, the total process time of 3 hours and 39 minutes is not unreasonable for production. However, in the sieve analysis, the coarsest fraction disappears early in the process, and well into spraying the fraction of material defined as "fines" continues to increase to nearly 20%. Ultimately, these fines are agglomerated to the main fraction of material, and their final content is marginal. However, the average particle size shifts to the coarse side, with more than 40% retained on the 40 mesh screen compared with only 18% found in the batch coated using the HS components.

The process time using the HS components was reduced by 48 minutes or approximately 30%. In addition, the near absence of fines seen in the SEMs will lead to improved batch-to-batch reproducibility in the dissolution profile.

35.8 SCALE-UP OF WURSTER PROCESSING

"Is the process scalable?" is a reasonable and common question. The answer typically begins "yes, but . . ." The ability to scale a process successfully depends on the magnitude of difference between the small and larger-sized batches. In the laboratory, development may take place with a batch size smaller than 1 kg. This size of equipment should be considered to be for screening purposes, without expecting that the formulation developed on the small scale will behave identically in pilot or production equipment. The batch size is small, the particle trajectory is comparatively short, the bed depth is shallow, and the spray nozzle does not use much compressed air to produce droplets. The mass of the batch is likely to be insufficient to manifest problems with attrition due to fragility of the substrate. Mass flow inside of the partition during spraying, cannot match that of larger equipment, so the amount of coating or layer applied to the substrate will not be in the same proportion as in larger equipment.

Experience has shown that the biggest hurdles in scale-up occur when moving from a 6″ Wurster coater (or smaller) to a 7″, 9″, 12″, or 18″ Wurster coater. Usable working volumes (courtesy of Glatt Air Techniques, Inc., Ramsey, NJ) are 3.6 L (6″ Wurster

coater), 8.3 L (7″), 13.6 L (9″), 37 L (12″), and 102 L (18″). In addition to the increase in volumetric capacity, the partition length increases, the partition diameter increases, bed depth and mass effects increase, particle trajectory increases, and the spray-nozzle dimensions increase (in size and consumption of atomizing air volume) to accommodate the increased spray rate. High kinetic energy from the combination of high atomizing air velocity and volume may be a significant, and in many cases unpredictable, factor in scale-up.

What should be evident is that scaling from any of the laboratory-sized Wurster systems to the pilot-scale 18″ Wurster coater is the true "scale-up," presenting the most significant challenges. Thereafter, production equipment typically incorporates multiples of the partitions and spray nozzles used in the 18″ Wurster. In reality, this "scale-up" should be referenced as "scale-out." Parameters such as process air volume, atomizing air pressure (and volume), spray rate, and temperatures derived for the 18″ Wurster are typically replicated on a "per partition" basis in the production equipment. Hardware components are also duplicated—the spray nozzle, partition diameter, and orifice plate configurations (permeability in both the up- and down-bed regions) are the same in both scales of equipment. Partition length is nominally longer, therefore the increase in bed depth is usually inconsequential provided that the 18″ Wurster coater is used at or near its working capacity. Examples of the pilot-scale 18″, and production scale 32″ HS Wurster coaters are shown in Figs. 35.2 and 35.25.

Successful scale-up in Wurster processing depends on a number of factors. At a minimum, development scientists should have a good understanding of the robustness of the process in laboratory equipment. Experiments bracketing the operating ranges of key parameters should be conducted, preferably on batch sizes exceeding 5 kg. Smaller batch sizes may be used, and the relationship of the process variables may be seen, but very small batches may not be indicative of what is seen on a somewhat larger scale. Ideally, design of experiments should be employed to quantify the magnitude of impact for the key variables.

FIGURE 35.25 Inside view of 32″ HS Wurster coater. *Source: Photograph courtesy from Glatt Air Techniques, Inc.*

As in any process, there is a multitude of variables to consider, including batch size, spray rate, atomizing air pressure/volume, process air volume, and temperatures (product, process air, and dew point). There are also equipment considerations—down-bed and up-bed orifice plate configuration, partition height above the plate, and the spray nozzle type, as mentioned previously.

Some concerns are not readily apparent. Equipment may increase in size and batch capacity, but not all components increase proportionally. For example, outlet air filter surface area actually decreases in proportion to batch size. This is problematic for products using layering or coating materials that cause filter media to blind. When a filter becomes occluded, process air volume is affected, particularly during shaking. The Wurster process requires process air flow to be continuous—it is not interrupted to shake fines back into the bed. As a consequence, when it is time to shake the filter, air flow is stopped, but only through the filter segment to be shaken. In contemporary machines, this means that the totality of the process air volume must now pass through the remaining half of the filter media. If the differential pressure across the filter is high, there will be a precipitous drop in process air volume. A consequence is that the velocity of substrate inside of the partition will momentarily drop, which could cause agglomeration. In an extreme case, with multiple partition production-scale Wurster systems, a severe drop in total process air volume may cause one of the partitions to stop "spouting." Under this circumstance, this partition may be filled from the top by the remaining spouting partitions. Even after shaking stops, and the process air volume returns to its set point, the filled partition may not evacuate. Atomizing air from the spray nozzle inside of the partition will create a small void, into which the liquid continues to be sprayed. However, the material surrounding this void is static, and is quickly over-wetted. Eventually, agglomerates and wet mass will make their way outside of the partition, and sampling will reveal a severe problem. By this time, the batch is essentially lost. Ultimately, if laboratory experimentation has shown that the spray material has a propensity to cause filter blinding, even to a comparatively limited extent, filter media must be a strong consideration during scale-up activities.

Relative to the environment in which spraying takes place (the coating zone), the particle size of the substrate is becoming smaller in scale-up. This may result

in a higher coating efficiency (coating material applied with respect to the theoretical or expected yield) in pilot and production equipment versus the laboratory scale. If the product is coated on a weight basis, the higher coating efficiency may lead to a change in dissolution (for sustained release coatings, it would decrease).

35.8.1 Batch size

As mentioned previously, the maximum batch size for a Wurster process is identified as that which occupies all of the volume outside of the partitions with them at rest on the orifice plate. There is some flexibility in batch size—the process will still work when using less than the maximum working capacity. However, it has been seen in scale-up that batch size, within a given insert size, may impact finished product properties. Therefore, it is recommended that the proposed finished batch size be close to the working capacity of the larger Wurster (approximately 75−100% based on the finished product density).

35.8.2 Spray rate

The spray rate for a product is typically a key variable, from several perspectives. The first is economic—long processes result in high manufacturing costs. Lengthy processes also increase the likelihood of problems during the process, particularly nozzle-port clogging (there are some hardware alterations that can be made to eliminate or at least mitigate this potential problem). Spray rate also dictates the rate of accumulation of solids by the batch, and this is important for finished dosage form performance. Coated product is nearly always stronger than uncoated core material. If the core sloughs fines, these may be incorporated into the layers of film, altering its release properties (release will be governed more by imperfection than by the intrinsic properties of the film). Consequently, it is a goal in scale-up to maximize the solids addition rate.

Drying capacity is a key component in scale-up. The starting point for estimating the spray rate is related to the increase in process air volume (at the same temperature and dew point as used in the laboratory scale), not the increased batch size. However, the rate-limiting factor may not be completely related to the increase in drying capacity offered by the larger Wurster system. Irrespective of the physical properties of the liquid, the application rate will more likely be strongly controlled by the interface between the accelerating core material, and the atomized droplets (coating zone). There may be significant benefits in using the spray nozzle surround described previously as HS Wurster equipment, which is not typically used or is of lesser benefit in small Wurster equipment. The surround prevents the flowing substrate from entering the spray pattern until it is more fully developed. This eliminates local over-wetting, preventing agglomeration in this region. A consequence is that scale-up in the Wurster process may not necessarily mean a significant increase in process time, which is a common occurrence in many types of processes.

35.8.3 Droplet size and nozzle considerations

Spray conditions represent a challenge for development personnel in scale-up. Small Wurster coaters tend to use the same size nozzle (Schlick 970 series), and as the product proceeds through small increments in batch size (from 6″ to 7″ or 9″ units with batch capacities ranging from approximately 1 to 10 kg), only atomization air pressure needs to be increased to accommodate the somewhat faster spray rate. However, when shifting to the 12″ or 18″ Wurster coater, the potential increase in spray rate forces operation of this small nozzle beyond its performance envelope, and a move to a larger nozzle, with a higher atomizing air volume, is mandated. Droplet size is principally related to the air-to-liquid mass ratio, and attempting to keep this in the same range in the scale-up efforts is recommended. However, there is a caution to its use. In some calculations, the air-to-liquid mass ratio may result in atomizing air pressure values that are outside of the nozzle's recommended operating range. The spray nozzle performs two different functions: (1) to produce droplets for the coating or layering application, and (2) to attempt to prevent the nozzle from fouling. The atomizing air at the tip of the nozzle helps to clear the tip of material forming as a result of drying that occurs between the interface of the liquid and the very dry atomizing air. If the atomizing air pressure (below approximately 1.5 bar) and volume is too low, the velocity of the atomizing air will be insufficient to remove dried film from the nozzle tip. This material could eventually grow to a size that impacts or diverts the spray pattern, causing a sudden and severe amount of agglomeration. In addition, the atomizing air velocity must be sufficient to permit shearing of the liquid to produce droplets. If the pressure and velocity are too low, agglomeration will be unavoidable.

At the other end of the spectrum, if the air-to-liquid mass ratio calculation yields an atomizing air pressure exceeding approximately 3.0 bar in the pilot-scale Wurster, attrition of the core material may be the consequence. Kinetic energy is represented as $E = 1/2\ mv^2$, where "m" is mass and "v" is velocity. Assume a scale-up from a 7″ Wurster machine (laboratory scale), with a batch size of approximately 5 kg, to a pilot-scale 18″ HS Wurster machine, with a batch size of approximately 60 kg. At 2.0 bar atomizing air pressure, the 970 series

(laboratory) nozzle uses about 2.1 cfm of compressed air. At the same pressure, the HS nozzle consumes nearly 30 cfm, or about 14 times that required for the laboratory batch. The atomizing air velocities are essentially the same—both operate at supersonic speeds at this pressure. If constant air density is assumed, the kinetic energy equation shows that the material in the pilot-scale equipment will be exposed to nearly 14 times more energy than the laboratory batch. For most applications using either water- or solvent-based materials, an atomizing air pressure of 1.8–3.0 bar is sufficient.

35.8.4 Process air volume

Process air volume is also a key factor in scale-up, providing three major functions. First, it delivers heat to the product, for evaporation and removal of the coating application medium. Second, and perhaps more significantly, it strongly affects the fluidization pattern. For scale-up, approximately the same air velocity through the partition plate for both sizes of inserts is recommended. The increase in overall air volume will then be principally related to the increase in the partition plate area through which the fluidization air will flow in the larger machine. "Scale-up" generally refers to increasing batch size and equipment geometry from small (6″, 7″, 9″, and 12″ Wurster coaters) to the pilot-scale. Partition diameter length increases, up to and including the 18″ Wurster coater. Beyond the 18″ Wurster coater the concept of "scale-out" may be more applicable, and scaling to 32″, 46″, or 55″ Wurster coaters is more direct. The larger inserts use multiples of the same diameter partition and spray nozzle found in the 18″ Wurster machine. Therefore, the increase in air flow would be a multiple of the number of partitions (700 cfm in the 18″ would lead to a target air flow of 2100 cfm in the three-partition 32″ Wurster machine).

The final manner in which the air volume impacts the process in scale-up is also related to the fluidization pattern. In multiple-partition Wurster systems (such as the 32″ and 46″ coaters), distribution of the process air across the entirety of the product container is a function of several factors. Orifice plates in both the up- and down-bed regions are selected based on the size, density, and surface properties of the materials being produced. As stated previously, the goal is to have the process air delivered to the product bed, such that the material in the down bed is in "near-weightless suspension" or behaving like a fluid. To achieve an essentially level surface requires a minimum of pressure across the down bed. This pressure is the result of several factors: core material quantity (initial batch size); partition height (the influence of the Venturi effect); orifice plate configuration; and finally, process air volume. For example, if 700 cfm were used in the 18″ Wurster coater, the target value for the three-partition 32″ Wurster coater would be 2100 cfm, as previously stated. Initial fluidization with the desired batch size may result in a tilted bed (higher in the rear of the insert than in the front) if product differential pressure was insufficient to cause the air to be uniformly distributed. Raising the partition height somewhat would increase the mass flow in the up-bed region, potentially increasing overall product differential pressure. If this change were insufficient, a second approach would be to use a less permeable down-bed plate. Both of these alternatives are preferred to increasing the process air volume. The consequent elevated spout height may enhance the potential for attrition, due to collision with the inside components of the filter housing.

For multiple partition production Wurster systems, it is prudent to optimize the fluidization pattern prior to actually producing a batch. A "mass flow study," which involves bracketing of process air volume and partition height for a set of orifice plates and batch size is recommended. The first step is to configure the Wurster insert with a set of up- and down-bed orifice plates. The initial partition height for the trial would be set at the lowest value selected for the study. Finally, the insert is positioned in the machine tower, and process air is drawn through it. The range of process air volumes to be tested should bracket the air volumes to be used for all steps of the process. At each of the selected air volumes, the product differential pressure (dP product) should be recorded. These will be the baseline contributions of the orifice plates. When the process is repeated with a batch in the insert, the difference between the total dP product and the baseline will reflect the "mass flow" in the insert. The goal is to identify the maximum product contribution, with a minimum of disruption to the fluidization pattern. An example is shown in Figs. 35.26 and 35.27. In Fig. 35.26, the air volumes tested are 1250, 1500, 1750, 2000, 2250, and 2500 cfm. At each increase in air volume, there is a slight increase in product differential.

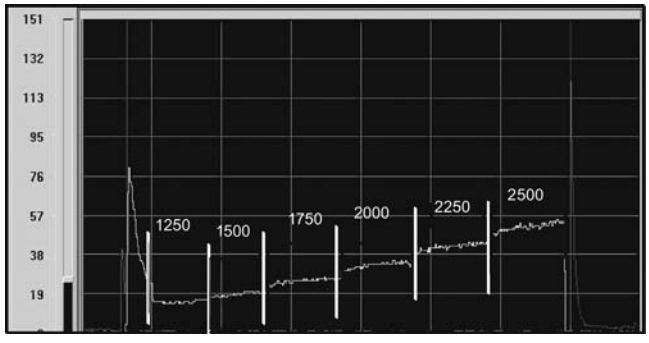


FIGURE 35.26 Product differential pressure in an empty 32″ HS Wurster coater fitted with type "B" and "G" orifice plates (air volumes displayed are in cfm).

Overall, the pressure contribution from the plates ranges from about 17 to 55 mmWC. Fig. 35.27 shows the impact of putting a large batch (approximately 250 kg) of pellets into the insert (note: it is prudent to use a very strong substrate for these trials because they will be subjected to mechanical stress for an extended period of time). With the partition height set at 50 mm, it is evident that as air volume is increased, the quantity of product in motion (mass flow) increases. The values for 1250 and 1500 cfm seem to indicate that the air volume is insufficient to fluidize the entire batch—for the bed is certainly stagnant in some region at 1250 cfm. At 1500–2000 cfm, the peak-to-trough values indicate a reasonably stable fluidization pattern. At 2250 and 2500 cfm, the enlarged peak-to-trough values reveal that some air bubble coalescence is occurring, resulting in some degree of back flow or turbulence in the down bed. This is not necessarily negative—a periodic bubble bursting through the down bed, along the wall of the product container, gives assurance that the product is not stagnant in this region. The air volume selected for this particular process was 2250 cfm. In addition to the fluidization properties, the high volume of air permits higher rates of heat and mass transfer, which translate to faster spray rates and a shorter process time.

Higher partition heights, common for pellet coating, were also tested (60 and 70 mm). Interestingly, the peak-to-trough values widen at lower air volumes with higher partition heights. It is speculated that the strength of the venturi at the base of the partition, which draws product into the up bed, is either lessened or defeated, allowing air bubbles to move only vertically in the down bed.

The previous example is effective for coating or layering materials that do not exhibit tackiness in the down bed—they have a low surface coefficient of friction. Unfortunately, there are many products that display tacky behavior, and these same processing conditions may permit stalling of the batch in some region. Some users may try to counter this behavior by increasing the overall process air volume. However, it is likely that the majority of this increase will flow through the partition, and this region is not the source of the problem. It would be desirable to divert more of the process air volume to the down-bed region. A method for achieving this would be to replace the up-bed plates with ones that are less permeable. In this manner, at the same air volume, more of the process air is diverted through the down bed, making it more vigorous or turbulent. Figs. 35.28 and 35.29 illustrate a mass flow study that is a consequence of this change. The "B" down-bed plate remains, but the "G" up-bed plates were replaced by the less permeable "H" plates. Once again, the partition height is set at 50 mm for the test. What can be seen is that the dP product contribution of the plates alone is

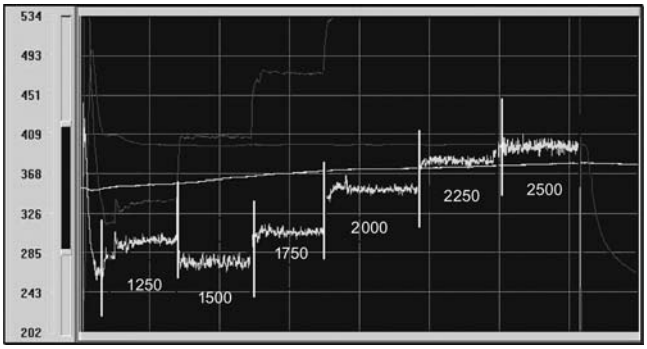


FIGURE 35.27 Product differential pressure in a 32″ HS Wurster coater fitted with type "B" and "G" orifice plates, with a batch (air volumes displayed are in cfm).

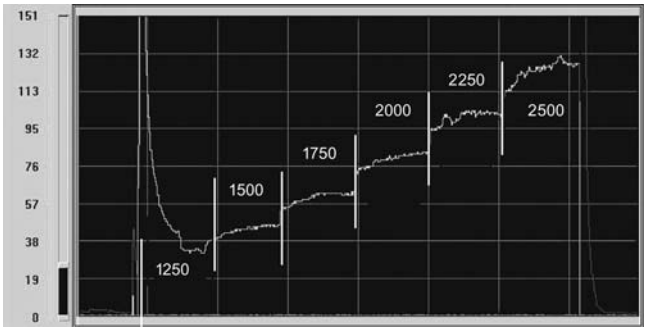


FIGURE 35.28 Product differential pressure in an empty 32″ HS Wurster coater fitted with type "B" and "H" orifice plates (air volumes displayed are in cfm).

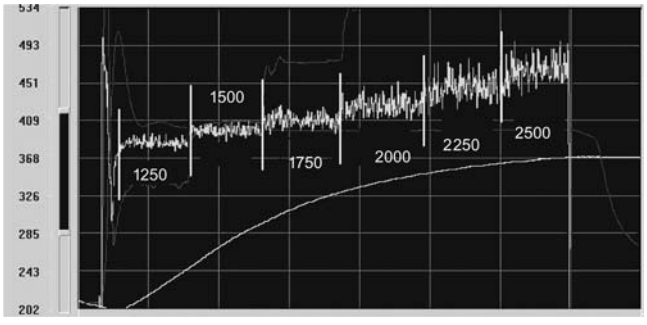


FIGURE 35.29 Product differential pressure in a 32″ HS Wurster coater fitted with type "B" and "H" orifice plates, with a batch (air volumes displayed are in cfm).

now considerably higher. In this case, the range is from approximately 30 to 125 mmWC. When the same batch weight is added, stable down-bed flow occurs at much lower air volumes—somewhere between 1250 and 1500 cfm. The peak-to-trough values for higher air volumes are very broad compared with the previous example, indicating significantly more turbulence in the down bed. The selected air volume for this trial was 2000 cfm. The turbulence was sufficient to avoid

regional bed stalling with the tacky coating material, and the relatively high air volume allowed for a reasonably high spray rate.

35.8.5 Process air and product temperatures

As is typical of the laboratory scale, the temperature of the process air is generally adjusted to maintain a constant product temperature, and as spray rate is increased, the process air temperature is adjusted upward. If laboratory trials were conducted using a conservative or low product temperature, it may be possible to explore a higher product temperature in the production-scale equipment to improve productivity. However, it is strongly recommended that this is not done if the impact of higher product temperature on finished product attributes has not been explored, particularly for latex-coating materials.

35.8.6 Mass effects

The influence of larger batch size (or mass effects, as they are referred to), is more difficult to predict. In small-scale equipment, the product must be robust—if it must be treated cautiously in laboratory equipment, scale-up to pilot and production machinery will be nearly impossible. In Wurster processing, the big step in scale-up is from laboratory equipment (6″, 7″, 9″) to the pilot-scale 18″ Wurster machine. In small machines, bed depth rarely exceeds 200 mm, and fluidization height (substrate travel out of the partition) is principally limited by machine dimensions to 1.25 m or less. In the 18″ unit, bed depth may range up to 600 mm, and fluidization spout height can approach or exceed 2 m. Product may also collide into the outlet air filters or other mechanical components of the machine. Further scale-up to the 32″ or larger Wurster coaters from the 18″ is somewhat simplified because bed depth is about the same, and the objective in fluidization is to keep particle velocity similar in larger equipment to that used in the pilot-scale machine.

35.9 SUMMARY

The Wurster system is widely used for layering and coating because of its ability to apply high-quality films to a broad range of substrates. The orifice plate configuration in combination with the partition height and process air volume, organize the substrate in close proximity to the spray nozzle. Droplets of coating liquid travel only a short distance, and are applied cocurrent with the accelerating substrate. Productivity is related to the presentation of substrate surface area per unit time with respect to this quantity of liquid. Layering and/or film coating may be conducted using solutions or suspensions of materials in liquids comprising organic solvents or water. The interaction between droplets and substrate is so rapid that materials applied using even the most volatile solvents are not subject to spray drying—they are high in quality, and their release is governed by their intrinsic properties, not by imperfections in the film. In addition, within limits, materials can be sprayed from a molten state, congealing onto the surface of the slightly cooler substrate. As a production-scale batch processor, the Wurster has a moderate batch capacity (up to approximately 600 kg), and is efficient in terms of material balance and productivity.

CHAPTER

36

Commercial Manufacturing and Product Quality

D.Y. Peng[1], *Y. Hu*[2], *S. Chatterjee*[3] *and D. Zhou*[2]

[1]Shire Pharmaceuticals Plc, Wayne, PA, United States [2]AbbVie, North Chicago, IL, United States [3]U.S. Food and Drug Administration, Silver Spring, MD, United States

36.1 INTRODUCTION

Over the past several decades, the pharmaceutical industry has grown tremendously, and significant progress has been seen in improving drug product quality and reliability of pharmaceutical manufacturing process. However, the public is still facing unprecedented drug shortages and recalls in recent years. By far, the most frequently cited reasons for drug shortages relate to manufacturing and quality issues.[1,2] In addition, the pharmaceutical industry is also facing several challenges such as payer's rising expectations on lowering cost while ensuring good product quality with a reliable supply, increasing the complexity of manufacturing technology, and increasing regulatory scrutiny. Hence, the industry is increasingly focusing on better understanding and control of commercial manufacturing as it relates to the product quality. A robust commercial manufacturing process means it is capable of consistently manufacturing desired quality product that meets the needs of the patient. A robust process development starts with a systematic approach to product and process design: understanding the impact of variability in input variables (raw materials, formulation and process parameters), and defining an appropriate control strategy to mitigate potential risks to product quality and manufacturability. Robust process development also includes a demonstration of the capability of the process at commercial scale and approaches to continually monitor and identify improvement opportunities for the process over the lifecycle of the product. These activities are the key three stages of the lifecycle approach for process validation as discussed in the 2011 Food and Drug Administration (FDA) process validation guidance titled *Guidance for Industry: Process Validation: General Principles and Practices*.[3] This guidance defines the process validation as the "collection and evaluation of data, from the process design stage through commercial production, which establishes scientific evidence that a process is capable of consistently delivering quality product." This definition of process validation focuses on scientific evidence throughout the product lifecycle. Process validation should not be considered as a one-off event in producing three consecutive batches of commercial scale product and evaluating its conformance with the predefined specifications. Instead, it is a science and risk-based development paradigm that builds product quality through a three-stage lifecycle approach which links the product and process development, qualification of the commercial manufacturing process, and maintenance of the process in a state of control during commercial production. In 2014, European Medicine Agency issued a revised guidance on process validation.[4] Even though there are differences in terminology, both guidances emphasize a lifecycle approach and embrace the basic principles of science and risk-based approach put forth in ICH Q8, Q9, and Q10. Since the introduction of these guidances, much progress has been made by the pharmaceutical community to address the expectations and challenges in implementation during process development and commercial manufacturing. This chapter will discuss some of the useful approaches that can be used to achieve a robust commercial manufacturing using the three-stage lifecycle approach, implementation of science and risk-based approaches such as advanced process analytical technology (PAT) tools, and real-time release testing

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00036-4

(RTRT) strategies. These concepts have the potential to provide an increased assurance of product quality.

36.2 PROCESS DESIGN, UNDERSTANDING, AND CONTROL STRATEGY DEVELOPMENT

The initial stage (Stage 1) of process validation is process design, understanding, and control strategy development. It is well recognized that a rational strategy for process development is to begin with the end in mind and to focus on patient's need with regards to drug product safety and efficacy. It is also important to consider the product manufacturability which is defined as the extent to which a product can be easily and effectively manufactured at minimum cost and with maximum reliability.[5] The first step of product and process development is to establish the quality target product profile (QTPP), which is a prospective summary of the quality characteristics of a drug product that ideally will be achieved to ensure the desired quality, taking into account safety and efficacy of the drug product.[6,7] QTPP forms the basis of design for the development of the product and process. From QTPP, the critical quality attributes (CQAs) of the drug product can be identified. CQA is defined as the physical, chemical, biological, or microbiological property or characteristic that should be within an appropriate limit, range, or distribution to ensure the desired product quality.[4] Criticality of a quality attribute is primarily based on the projected severity of harm to the patient (eg, from loss of drug efficacy or causing serious adverse effect) should the product fall outside the acceptable range for that attribute. The probability of occurrence, detectability, and controllability are not relevant to the determination of the criticality of a product quality attribute.[5]

Once the QTPP and CQAs are defined and identified, the subsequent development activities focus on how to achieve the QTPP design requirements. Risk assessment tools are often used to prioritize development studies.[8] Prior knowledge of existing commercial manufacturing of similar product or manufacturing technology platform can be utilized to facilitate the risk assessment. Based on the initial risk identification, potentially high-risk material attributes, formulation variables, and process parameters are further studied to understand their impact on the identified CQAs. Based on the development study results and updated risk analysis, critical material attributes (CMAs) and critical process parameters (CPPs) can be identified, and the functional relationships between CMAs, CPPs, and drug product CQAs can then be established during development. Knowledge gained through appropriately designed development studies culminates in the establishment of a control strategy. A control strategy is a planned set of controls, derived from current product and process understanding that ensures process performance and product quality. The controls can include parameters and attributes related to drug substance and drug product materials and components, facility and equipment operating conditions, in-process controls, finished product specifications, and the associated methods and frequency of monitoring and control.[9] Yu, et al. summarizes the commonly used control strategy in three levels (Fig. 36.1).[10]

Level-3 control strategy mainly relies on extensive end-product testing and tightly constrained material attributes and process parameters. This type of control strategy leads to acceptable product quality when the input variables are controlled adequately. However, when the inputs vary beyond expected variability, the resulting product quality can be impacted. The pharmaceutical industry has recognized the limitation of the Level-3 control strategy and has been developing more advanced control strategies in the past years.

Level-2 control strategy consists of controls with reduced end-product testing and allowing variation of material attributes and process parameters within the established design space. Understanding the impact that variability has on in-process materials, downstream processing, and drug product quality provides an opportunity to shift controls upstream and to reduce the reliance on end-product testing. Among process monitoring technologies, PAT has gained attention in pharmaceutical industry since the publication of PAT guidance in September 2004.[11] In this guidance, PAT is defined as "A system for designing, analyzing, and controlling

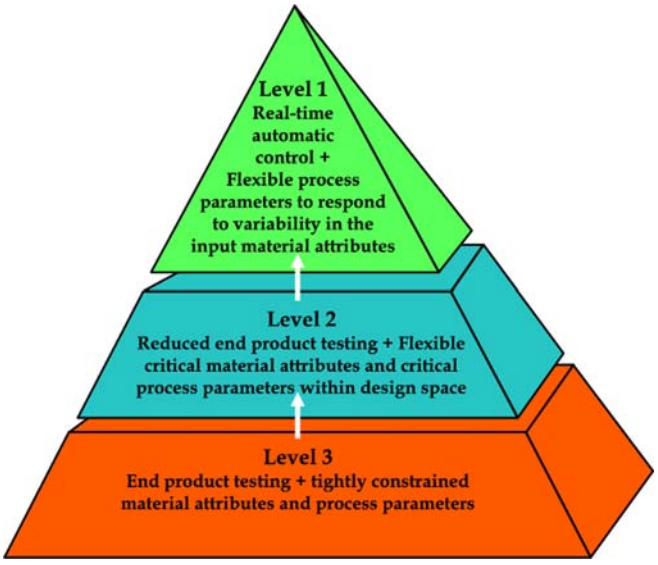


FIGURE 36.1 Control strategy implementation options. *Copied from Yu LX, Amidon G, Khan MA, Hoag SW, Polli J, Raju GK, et al. Understanding pharmaceutical quality by design. AAPS J 2014;**16**:771–83.*

manufacturing through timely measurements (ie, during processing) of critical quality and performance attributes of raw and in-process materials and processes, with the goal of ensuring final product quality." The PAT toolkit often includes process analyzers, multivariate analysis tools, process control tools, and continual improvement/knowledge management systems.

A recently published review paper on PAT discussed the advantages of PAT implementation at various stages of pharmaceutical development, and recent developments in sensor technology, followed by specific cases of PAT implementation in laboratory and pharmaceutical production.[12] Additionally, RTRT could be regarded an example of Level-2 control, as it provides the ability to evaluate and ensure the quality of final product based on in-process data, which typically include a valid combination of measured material attributes and process controls. RTRT is defined as "the ability to evaluate and ensure the acceptable quality of in-process and/or final product based on process data."[6]

Level-1 control strategy utilizes automatic engineering control to monitor the CQAs of the output materials in real time to achieve adaptive control. The key concept is to monitor input material attributes and automatically adjust the process parameters based on the enhanced process understanding to assure that CQAs consistently comply with the established acceptance criteria.[10] This includes control strategy options such as feedback or feed forward control. These are examples of dynamic control where downstream process parameters and attributes are manipulated in response to variation in upstream process parameters and attributes. A feedforward control is defined as one in which manipulations are made proactively to downstream variables in response to measurements of disturbances upstream, whereas feedback control is one where manipulations are made to upstream variables response to measurement of disturbances downstream. On-line monitoring techniques such as PAT could facilitate implementation of such dynamic control strategies.[6] It is noteworthy to mention that PAT tools and RTRT are not necessarily always implemented together. A PAT tool can be implemented for the monitoring of attributes of either in-coming materials or that of in-process materials during manufacturing. PAT tools can also be put in place for monitoring quality of finished products, without using that information for product release. On the other hand, adoption of PAT is not the only way to implement RTRT. For example, RTRT can be achieved by the use of predictive models as a surrogate for traditional release test, where the model uses traditional in-process measurements. (This is often referred to as "soft sensors.")

This chapter only provides a high-level overview of process design, understanding, and control strategy development. More in-depth discussion of each unit operation and its unique challenges and solutions can be found in other chapters of this book and related scientific literature.[6–8,13–22]

36.3 PROCESS SCALE-UP, TECHNOLOGY TRANSFER, AND PROCESS QUALIFICATION

Process scale-up and technology transfer in pharmaceutical industry involve, in general, moving a product/process from research and development stage into the production stage. More detailed discussions can be found in chapter "Product and Process Development of Solid Oral Dosage Forms," and other chapters on process development and scale-up of individual unit operations. In most cases, early development studies are often conducted using laboratory or pilot scale equipment. The potential risks associated with scaling-up the process from lab or pilot scale to production scale are assessed and mitigated via further studies. The assessment includes evaluating the impact of facility, utility, environment conditions, personnel and equipment design and operating principle changes on product quality and process performance. In some cases, a process step for example milling or compression is continuous operation in nature, the scale up can be achieved by simply extended the processing time. In some cases, the performance of a process can be predicted by fundamental engineering and scientific principles or empirical process models developed at the laboratory or pilot scale; the scale up is guided by the model prediction. While in other cases, at-scale knowledge is needed to adjust the control strategy developed at the laboratory or pilot scale. It is important to assess systematically product and process knowledge gained during the development stage and determine any residual risks within the proposed control strategy. At the end of Stage 1 (development phase), the proposed commercial control strategy is communicated to the receiving sites along with any identified residual risks which will be evaluated during the process qualification stage (ie, Stage 2 of the lifecycle approach for process validation). Several challenges regarding process qualification are discussed in the subsequent sections (36.3.1–36.3.3).

36.3.1 Design of a facility and qualification of utilities and equipment

The objective of process qualification is to evaluate the process designed and developed from Stage 1 (the initial stage of the lifecycle approach for process validation) and to determine if it is capable of

reproducible commercial manufacturing. To meet this objective, the first step is to select a suitable facility which is appropriately designed for the selected manufacturing process. For example, the facility design and environment control for an aseptic filling process for an injectable product would be considerably different from a solid oral dosage form manufacturing facility. The second step is to demonstrate that utilities and equipment are suitable for their intended use and perform properly. The qualification will include all the aspects that would traditionally fall into the installation qualification, operational qualification, and performance qualification (IQ/OQ/PQ) categorization. It is also important to ensure that the analytical methods for in-process controls and release testing are appropriately verified or validated for its intended use. Further detail information on these topics can be found in scientific literature and general guidelines developed by industry working groups.[23–26] These upfront preparation works are important to ensure the variability from the facility, utilities, and equipment, and the analytical methods are understood and controlled appropriately to support the process performance qualification (PPQ) execution.

36.3.2 Number of PPQ batches

Since the publication of the Process Validation Guidance, an often debated topic is how many PPQ batches are needed at Stage 2b to demonstrate the commercial manufacturing process performs as expected and the product manufactured at production site is ready for commercial distribution. There is no sacrosanct number of batches or approach for demonstrating a successful PPQ. Manufacturers are responsible for providing assurance that the process is adequately qualified before moving to routine commercial manufacturing in Validation Stage 3. Use of science and risk-based approach including statistical method to determine the number of PPQ batches is extensively discussed in scientific literature. Such risk-based approach combines knowledge of the capability of the process during development, fundamental scientific understanding of a particular process as well as the plan for future monitoring in Stage 3. Several example approaches reported in scientific literature are given here. For example, F. Wiles proposed using a combination of risk assessment, control charting, and capability statistics to determine how many PPQ batches are needed and when PPQ milestones have been met.[27] In this method, "an assessment of process risk is performed through application of a process failure mode, effects, and criticality analysis (PFMECA). The output of PFMECA is used to select appropriate levels of statistical confidence and coverage which, in turn, are used in capability calculations to determine when significant Stage 2 (PPQ) milestones have been met."[27] Success from the PPQ stage signals the readiness to release batches for commercial distribution, and the frequency of monitoring and testing can be adjusted during the routine commercial production stage. Individual control chart, moving range, and range/sigma control charts are used in conjunction with capability statistics to demonstrate that the commercial process is operating in a state of statistical control.

A Bayesian method was developed by H. Yang to address the question *"how many batches are needed for process validation under the new FDA Guidance."*[28] "By combining process knowledge gained from development stage with the expected outcomes of PPQ, the Bayesian method estimates the number of qualification batches that are needed to provide a high level of assurance that the process will consistently produce future batches meeting quality standards."[28] Several examples based on simulated data are presented to illustrate the use of the Bayesian method in helping manufacturers make risk-based decisions for Stage 2 PPQ. For example, a large "n" (=480) is needed for one process which has a poor history of performance during development, and small "n" (=4) for another process which shows a track record of good performance during development, and moderately large "n" (=18) for the third process which demonstrates variable performance during development.

The International Society for Pharmaceutical Engineering (ISPE) published discussion paper on potential approaches for answering the question, "How many PPQ batches are needed to demonstrate a high degree of assurance in the manufacturing process and that the control strategy is sufficiently robust to support commercial release of the product?"[29] This discussion paper describes a framework for systematically assessing the level of product knowledge and process understanding, and how well the control strategies are linked to the CQAs. Based on the residual risk identified from the risk assessment, three different approaches were proposed to estimate the number of PPQ batches and to establish the scientific evidence whether the level of between-batch variability is appropriate for commercialization:

1. Approach 1—based on scientific rationale and experience

   The paper proposed a range of number of PPQ batches based on the residual risk level. While 3 consecutive batches is sufficient for low-risk processes; more batches are needed when there is an elevated residual risk.
2. Approach 2—target product performance and target process capability

   This approach utilizes process capability index (C_{pk}) and target confidence level to determine the

number of PPQ batches based on the identified residual risk level. This is to address two questions: What is the objective measure that process will consistently produce a product that meet its requirement? and What is an acceptable high degree of assurance? The paper encouraged the manufacturer to justify the choices of the target confidence level and capability level based on the risk associated with product knowledge, understanding, and robustness of the control strategy.

3. Approach 3—expected coverage

   This approach is based on the concept of order statistics.[30] In this approach, the residual risk level is used to determine the expected "coverage" of the PPQ batches and the number of PPQ batches. The "coverage" is defined as the proportion of the expected "future" data to be contained within the acceptance limits. The lower the risk level, the lower the expected coverage from the PPQ batches alone, and the lower the number of PPQ batches needed.

Another frequently debated topic is how to select the appropriate conditions for PPQ batches. The concept of worst-case conditions for process validation was a key theme of the 1987 process validation guidance. However, attempting to cover worst-case conditions in process validation would often mean that parameters applied to validation batches bore a little resemblance to the actual variability of the process to be encountered during routine commercial manufacturing. Following the process design stage (Stage 1) where the process variability and the functional relationships between CMAs, CPPs, and the drug product CQAs are well understood and addressed, the PPQ batches should be manufactured under normal conditions by the personnel routinely expected to perform each step of each unit operation in the process. A risk-based approach is highly encouraged to select the conditions for the PPQ batches and justify it based on the enhanced understanding of the product and process and the actual variability of the CMAs and CPPs.[31] Thus, high assurance is obtained that the commercial manufacturing process is capable of consistently producing acceptable quality products within commercial manufacturing conditions.

36.3.3 Heightened level of monitoring and testing to demonstrate intra- and inter-batch consistency

As indicted in the 2011 FDA PV guidance, in most cases, PPQ will have a higher level of sampling, additional testing, and greater scrutiny of process performance than would be typical of routine commercial production.[3] The objective of the heightened level of monitoring and testing is to confirm uniform product quality throughout the batch or batches. The selection of the parameters to be monitored or tested is mainly based on their impact to product safety, efficacy and quality. In general, it can include, but is not limited to, the drug product CQAs, input material CMAs, CPPs, critical in-process controls, and other relevant parameters that are indicative of process performance. The increased level of the scrutiny and the duration of the heightened sampling and monitoring period can be adjusted based on the volume of production, process complexity, level of process understanding, and experience with similar products and process.

Yu et al. proposed a staged sampling approach when limited batches have been manufactured during the PPQ stage.[32] Fig. 36.2 represents an illustrated example using simulated assay data. The first three batches (A, B, and C), manufactured during the PPQ stage, represent a higher level of sampling within each batch. The goal of the initial higher level of sampling is to demonstrate product quality throughout each batch, establish initial estimates of within and between-batch variability, and use those estimates to generate an initial estimate of process performance. Then, sampling in subsequent batches (D, E, and F) is adjusted (in this case, lowered) to a statistically representative level that was based upon the variability estimates established in the first three batches.

Another example approach proposed by ISO 16269-6 (2014) uses statistical tolerance intervals.[33] In this approach, the coverage, defined as the proportion of the expected "future" data to be contained within the acceptance limits, is adjusted based on the risk of the product CQA. The actual number of samples required varies based on the data type (discrete or continuous) and the expected distribution (normal or unknown).[33]

Several other statistic approaches have also been proposed in a recent ISPE discussion paper including graphical exploration, variance components analysis, Monte Carlo simulation, ASTM E2709/E2810 (historically also referred to as CuDAL), statistical interval, Bayesian statistics, control charts, and process capability indices.[34] Even though this paper does not reflect a consensus position of the industry, it provides a comprehensive summary of viable approaches and options that can be used to evaluate the intrabatch and interbatch variability during process qualification stage.

When PAT and/or RTRT are used as a part of the commercial control strategy, special considerations should be given to the sampling strategy. Typically, implementation of RTRT would include in-process on-line/at-line sampling. The selected sample size is representative of the batch size and is also justified statistically. Additionally, the sampling frequency is chosen to provide assurance of consistent manufacture of

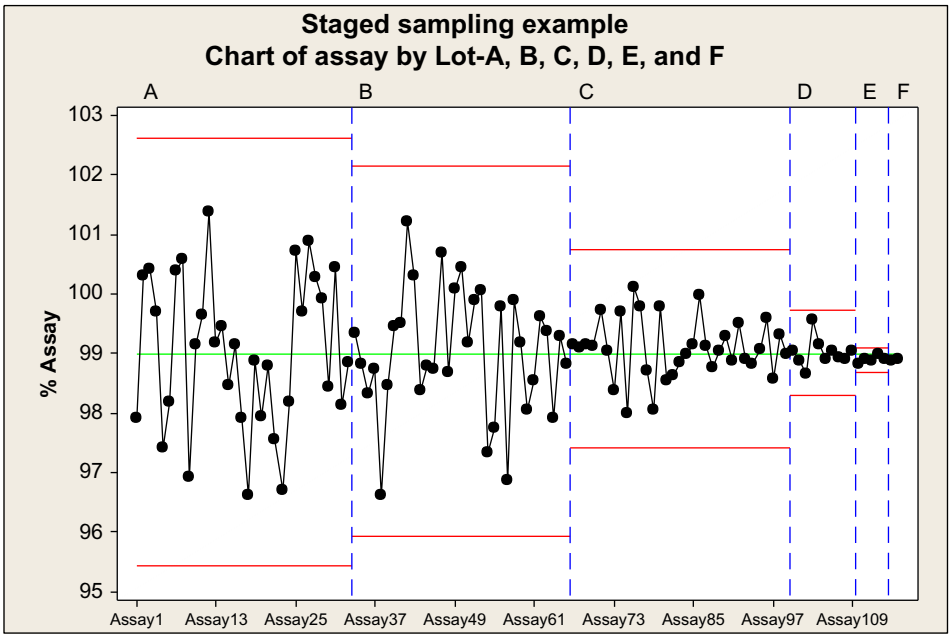


FIGURE 36.2 Theoretical example of a staged sampling approach when limited batches have been manufactured during process performance qualification (PPQ) stage.

desired quality product. In addition to sample size and sampling frequency, measurement location needs to be defined. For example, for real-time in-process identity testing for release, it is important to consider what/if any downstream processing steps are present after identification testing before final release of the finished product to prevent the inadvertent addition of any unknown materials into the process due to a system and/or human error. Furthermore, contingency plans can be in place in case of failure of the primary analytical method that is used for RTRT measurement. For example, if an NIR analytical procedure is used for tablet content uniformity to support RTRT, an alternate control strategy can be in place for an alternative mode of measurement of tablet content uniformity in case of sudden failures of the NIR instruments during routine commercial production.

36.4 CONTINUED PROCESS VERIFICATION

Often, significant sources of the variability related to commercial manufacturing have been systematically identified, and the appropriate detection, control, and/or mitigation strategies have been established during process design, scale-up technology transfer, and process qualification studies. However, during routine commercial manufacturing, a process may encounter additional sources of variation that were not previously detected or to which the process was not previously exposed. Hence, it is important to monitor the process continually and maintain the process in a state of control (the validated state) during commercial manufacture. The heightened testing and monitoring of process parameters, material attributes, and drug product quality attributes utilized during the PPQ can be continued until sufficient data are available to generate significant variability estimates. Testing and monitoring can then be adjusted to a statistically appropriate and representative level. For some cases, it is even necessary to requalify the process when the source of materials, equipment, production environment, personnel, and manufacturing procedures has been significantly changed.

36.4.1 Process monitoring program

In general, a science and risk-based approach can be applied to determine the process monitoring program, that is, what to monitor and the monitoring frequency and locations. A routine commercial manufacturing process has many input variables, process parameters, and in-process material attributes. The monitoring program focuses on the important parameters known or suspected to affect the quality, efficacy, or safety of the product. Additionally, other parameters can also be included if they are indicative of process performance and/or can provide a timely signal to assist in process controlling. Sources of information from other quality and manufacturing systems such as change control, scheduled preventive maintenance, calibration, or production interruptions; and customer complaints are also valuable inputs for monitoring. These may help explain unexpected shifts or variation in the production process and manufacturability issues.

In a recent publication, Mitchell et al.[35] provided an example of the monitoring and the review frequency for a process based on the enhanced understanding of the product and process from Stages 1 and 2 studies. In this example, high-risk parameters are monitored and reviewed more often to respond more quickly to out-of-control conditions that could risk product quality. Some medium risk process parameters and performance attributes are also included in the monitoring program since it is indicative of process performance and manufacturability. Low-risk parameters are not monitored due to their low impact on quality attributes. This example illustrates the principles that how the current understanding of the manufacturing process and the level of risk impacting product quality can be utilized to determine the parameters to be included in the monitoring program, the location and the frequency of the analysis.

36.4.2 Tools for process monitoring

Various statistical process control (SPC) tools can be used in monitoring and evaluating process variability and process capability. There are many available industry standards, books, and guides on these topics.[36–40] In the next section, two commonly used tools (control charts and process capability analysis) are briefly discussed.

36.4.2.1 Using control chart to evaluate if the process is in-control

Control charts were developed by Walter Shewhart in the early 1920s and are still widely used in many industry sectors including the pharmaceutical industry.[41–44] A control chart is a graphical display of a product quality characteristic that has been measured or computed periodically from a process at a defined frequency. The observations sampled at a particular time point constitute a subgroup. Through statistical methods, control charts can differentiate the common cause and special cause variation of a process. The common cause variation is natural or inherent to the process; while the special cause variation is not inherent to the process and may only be present at times. A process displaying only common cause variation is considered in a state of statistical control or stable while a process displaying special cause variation is not in statistical control. An illustrative example of a control chart can be seen in Fig. 36.3. In this figure, the process is monitored at a predetermined frequency, and the process output is measured and graphed sequentially. The process mean and standard deviation can be calculated. The upper SPC limit (UCL) and the lower SPC limit (LCL) are typically defined as 3 times of the standard deviation away from the process mean. The UCL and LCL establish an envelope around the process mean, where the majority of the output of a stable process should lie within. The control charts are sometimes also referred to as "the voice of the process."

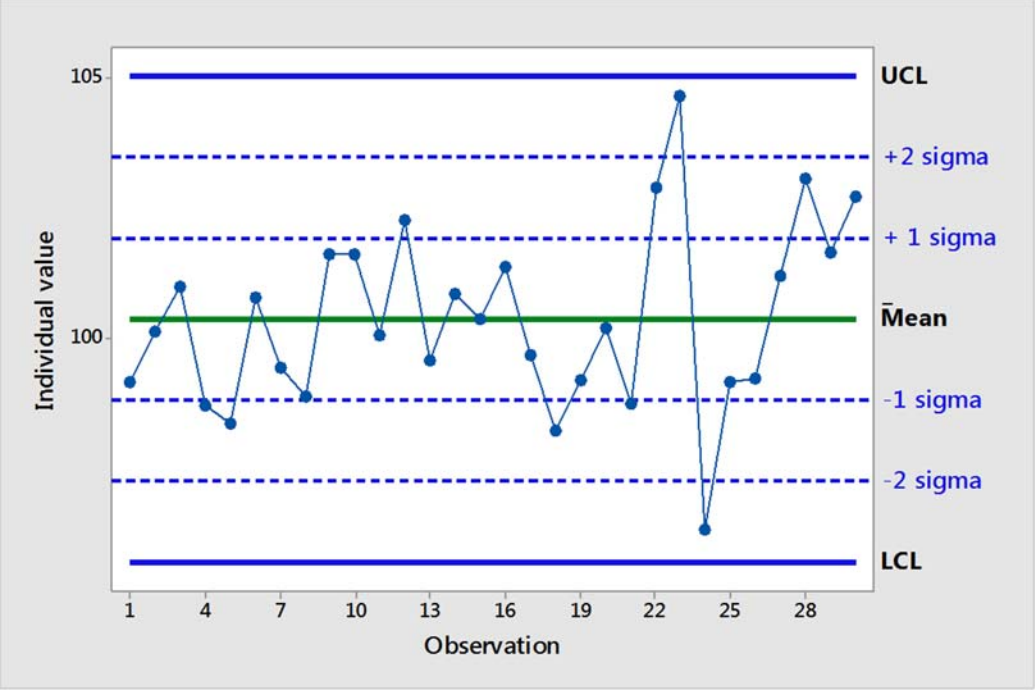


FIGURE 36.3 Schematic example of an I-chart.

To aid in the detection of special cause variations in the process, a set of decision rules are widely used along with the control charts. The set of decision rules is called the "Western Electric Rules," which was developed by the Western Electric Company, as the name implies.[45] The Western Electric Rules provided a standard way of identifying the out-of-control signals, by differentiating common cause variations from special cause variations. One frequently used Western Electric rule is "any single data point falls outside of the LCL and UCL" (the SPC limits). When a data point violates this rule, it indicates the process may have encountered special cause variations that may require further investigation. However, it is important to keep in mind the rule has a certain probability of being a "false alarm"; that is a data point from a stable process violates this rule just by chance. For example, for a normally distributed process, the probability of "false alarm" is close to 0.3%. Details of other Western Electric Rules can be found in the references.[33] It is important to establish a risk-based action plan that is based on the potential impact to product quality, when responding to the statistical signals from the control charts. Such a plan provides balance between over reaction to the signals and failure to act against unwanted variability that may impact product quality.[3]

There exist many different types of control charts, based on the characteristics of the process, the collected data set, and the underlying distribution assumptions. Fig. 36.3 is an example of a Shewhart control chart with individual measurement, called I-chart (subgroup size = 1). When subgroup size is between two and ten, an X-bar/R chart is most commonly used. When subgroup size is greater than ten, X-bar/S (standard deviation) chart becomes a more appropriate option. One key consideration of this type of charts is the rational subgrouping since the general assumption is that data in a subgroup are considered from a stable process to be used to define common cause variation. I-chart, Moving Range chart (MR-chart), X-bar/R chart and X-bar/S Chart are for continuous numeric variables. On the other hand, p-chart for the binomial distributed attributes (categorical or discrete numeric data, eg, by counting), and the u-chart for the Poisson distributed attributes are frequently used for discrete variables.[39]

These aforementioned univariate SPC charts works well if all the monitoring parameters are independent of each other. However, most pharmaceutical manufacturing process involves a large number of input material attributes, process parameters, intermediate and finished drug product CQAs, and some of these variables may be interrelated. Multivariate statistical process control (MSPC) can be a useful tool to address this challenge. MSPC refers to a set of advanced techniques for the monitoring and control of the manufacturing process (batch or continuous processes) which identifies the vital few variables (principle components) and the underlying patterns in a dataset. In addition, the functional relationships between these variables can be established (http://www.camo.com/applications/process-control.html). Furthermore, MSPC techniques can reduce the information contained within all of these variables down to two or three composite metrics through the application of principle components analysis or partial least squares. The use of MSPC in pharmaceutical industry has increased greatly in recent years.[46–49] The readers are encouraged to find further details of MSPC in the provided literature as it is beyond the scope of this chapter.

36.4.2.2 *Using process capability and process performance indices to evaluate if the process is capable*

While the control charts signify the "voice of the process," the specification limits (acceptance criteria) of the product quality attributes indicate the "voice of the customers" (based on patient's safety and efficacy needs). These two measures are compared to demonstrate if the process is capable of producing products meeting customer's requirements consistently. The comparison is often expressed in the form of a process capability index (C_p and C_{pk}). The calculation formula and description of these indices are presented in Table 36.1. The potential application of these process capability indices to enhance pharmaceutical product

TABLE 36.1 Description and Calculation of Process Capability Indices

Calculation	Description
$PC = 6\,\hat{\sigma}$	Process capability (PC): a statistical estimate of the outcome of a characteristic from a process that has been demonstrated to be in a state of statistical control
$C_p = \frac{(USL - LSL)}{6\,\hat{\sigma}}$	Process capability index (C_p): an index describing process capability in relation to specified tolerance
$C_{pkl} = \frac{Mean - LSL}{3\hat{\sigma}}$	Lower process capability index (C_{pkl}): an index describing process capability in relation to the lower specification limit
$C_{pku} = \frac{USL - Mean}{3\hat{\sigma}}$	Upper process capability index (C_{pku}): an index describing process capability in relation to the upper specification limit
$C_{pk} = \min(C_{pkl}, C_{pku})$	Minimum process capability index (C_{pk}): smaller of the upper process capability index and the lower process capability index

quality has been gaining international attentions in recent years.[32,50,51]

USL, upper specification limit; LSL, lower specification limit; $\hat{\sigma}$ represents common cause variability in a stable process

The process capability indices use the common cause variation as the variability measurement of the process. Therefore, it represents the "capability" of a stable process free of special cause variation. For a process that exhibits some special cause variations, more appropriate indices are P_p or P_{pk}, sometimes referred to as the process performance indices; where the process standard deviation estimate uses the overall standard deviation which accounts for all common cause variations and special cause variations.

It is important to point out that both the process capability and performance indices are statistical estimates of the process, and are subject to the variability of the sampling procedure. As such, the indices are less robust when the sample size is small; and become more reliable with increasing sample size. In fact, it is highly recommended that when a capability index or a performance index is reported, the lower bound of the estimate is also reported at the same time. For the process performance index P_{pk}, the approximate 100 $(1-\alpha)$% lower confidence bound of the process performance index[36] is defined as Eq. (36.1):

$$P_{pk} \geq \hat{P}_{pk} - z_{1-\alpha}\sqrt{\frac{1}{9n} + \frac{\hat{P}_{pk}^2}{2n-2}} \qquad (36.1)$$

where n is the sample size for the index estimate, and $Z_{1-\alpha}$ is the $100(1-\alpha)$ percentile of a standard normal distribution, and $\hat{P}_{pk}$ is the point estimate of the process performance index.

Fig. 36.4 shows the relationship between the lower boundary of P_{pk} and the sample size and confidence. For example, when a process demonstrates a process performance index of 1.15 with 60 data points, we can state such a process has a P_{pk} of 1.0 with about 90% confidence.

36.4.3 Continual improvement

Given the current regulatory requirement for reporting of postapproval changes according to 21 CFR 314.70, the pharmaceutical industry is dis-incentivized from making changes to an approved manufacturing process. For years, traditional pharmaceutical process development led to tightly controlled process and stringent specifications that were largely established on the basis of limited manufacturing experience. Product quality is ensured predominantly by restricting the flexibility of the manufacturing process and relying on extensive end product testing. Even though this approach has historically provided sufficient confidence that quality is suitably controlled in an unchanged process, this type of lifecycle management does not support continued assurance of quality and in the long term. It is neither efficient nor cost-effective and can result in reactive actions for both industry and regulatory authorities. There is a great need to embark the principle of "quality

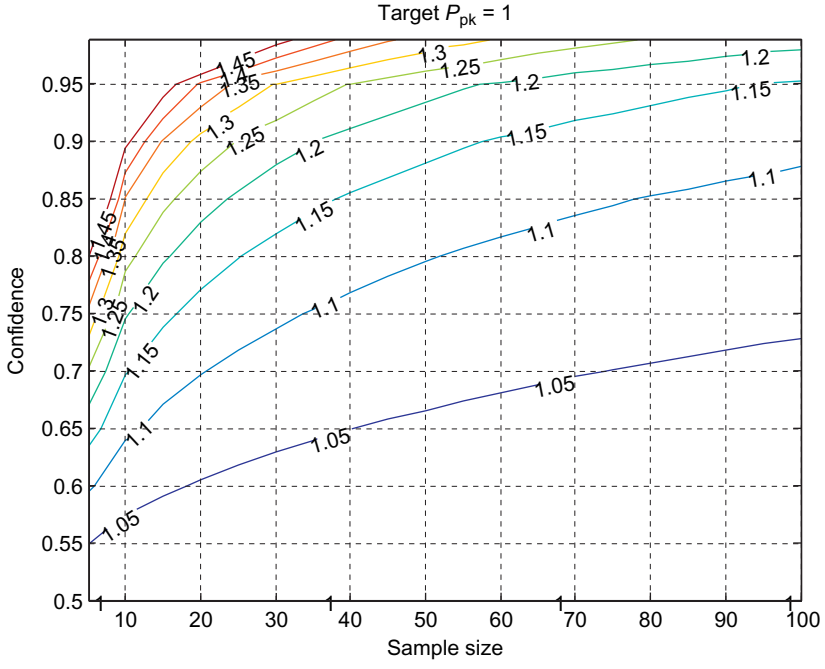


FIGURE 36.4 P_{pk} lower boundary.

excellence" whereby the pharmaceutical manufacturers and the regulatory agencies for collaborative partnerships to facilitate continual improvement efforts based on science and quality risk management.

Continual improvement is a set of activities that the manufacturer carries out to enhance its ability to achieve the desired product performance, and/or to make the process more efficient by increasing yield/reducing cost. Continual improvements typically consist of five phases- Define, Measure, Analyze, Improve, and Control (DMAIC)[52]:

- Define the problem and the specific project goals
- Measure the performance of the current process and collect relevant data
- Analyze the data to investigate and determine the cause-and-effect relationships. Seek out the root cause of the defect if any.
- Improve or optimize the current process based upon data analysis using techniques such as design experiments (DOE) to create a new state of the process.
- Control the newly established process to confirm the improvements. Implement control systems such as SPC, production dashboards, visual aids, and continued process verification (CPV) /monitoring.

36.4.4 Illustrative Example-1: monitoring key excipient material variability of an extended release tablets and continual improvement of process performance and product quality

A case study is given in this section as an example of the commercial manufacturing monitoring and continual improvement.[53] Niacin Extended Release (ER) Tablets is a hydrophilic matrix polymer based ER tablet which contains about 16% Hypromellose (HPMC) and 81% of API. A control chart (Fig. 36.5)[53] of niacin release at 20 hours (D20h) from more than 800 commercial lots indicated that the drug release was impacted by the HPMC lots, even though some variability was also apparent within the same HPMC lot, possibly due to additional variability in the manufacturing process and analytical method. Some of the tablet lots barely passed the lower specification limit of not less than 75% and a few lots failed. The apparent variability of the drug release at 20 hours of this product presented a significant challenge to maintain uninterrupted supply to patients of desired quality product.

HPMC is a cellulose derivative by introducing methoxy (MeO) and hydroxypropyl (HP) substitutions simultaneously under alkaline conditions with inherent chemical heterogeneity due to the potential chemical variability of cellulose derivatives. Initial investigation performed on HPMC lots used in the commercial manufacturing indicated that the HP content is a primary impacting factor ($p < 0.0001$) while the viscosity is of borderline significance ($p = 0.07$). Neither MeO nor particle size was found significant. A regression model based on the HP content and viscosity ($R^2 = 0.79$) predicts the D20h well in general. However, some HPMC lots deviated significantly from this model. As a part of the investigation, HPMC lots from two additional vendors (vendor B and vendor C) were also evaluated experimentally. The regression model did not predict dissolution for HPMC lots from vendor C (HP ~8.3–8.6%), although the predicted

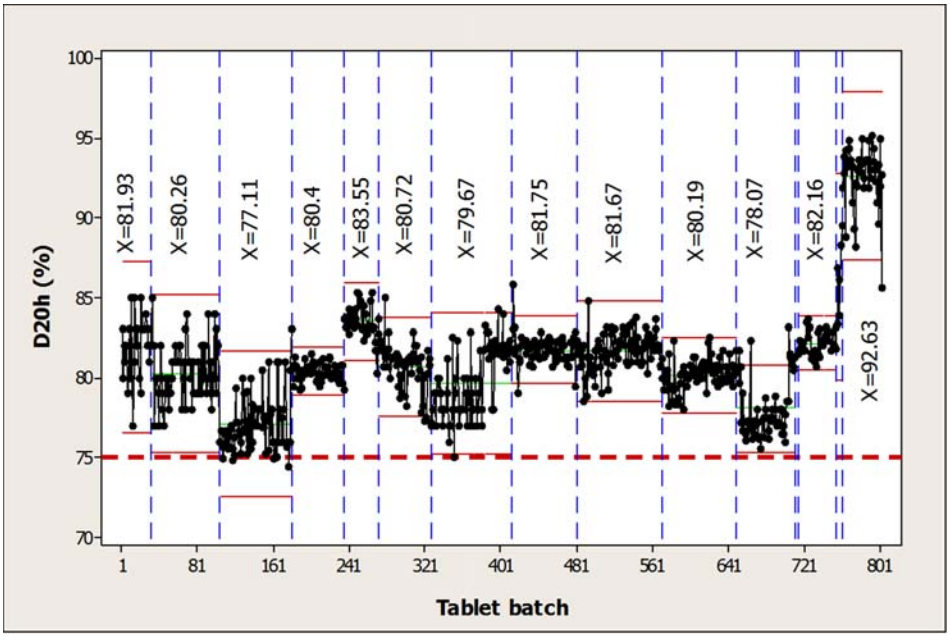


FIGURE 36.5 Run chart of niacin release at 20 hours segmented by HPMC lots.

and tested values were consistent (~90% released at 20 hours) for HPMC from vendor B (HP ~9.5–9.7%).

An investigation using a two-dimensional NMR clearly differentiated two types of hydroxypropyl substituents: one as native where the hydroxyl is free (designated as HP1) and the other is inner substitution where the hydroxyl group is further substituted by either a MeO group or another HP group (designated as HP2). The HP1/HP2 ratios were found to vary significantly among the vendors: HP1:HP2 ~ 3.3:1 for the commercial HPMC vendor, while HP1:HP2 ~ 2:1 for both vendor B and vendor C. The HP1:HP2 ratio is perceived to be important because HP2 represents a longer and larger substitution and greatly weakens the polymer/polymer interaction, thus, favoring faster drug release. Indeed, even though the overall HP contents for vendor C HPMC lots were much lower (HP ~8.3–8.6%), they produced fastest release rates (>90%) while the regression model predictions were low (<80%). It was postulated that the apparent conformance of vendor B materials to the regression model was purely a coincidence because with this highly substituted HP make up (HP1:HP2 ~ 2:1), the HP content does not need to be as high as in HPMC from the commercial vendor to release fast.

Even though the vendor-to-vendor differences were explained by the difference in the HP makeups, the fact that not all HPMC lots from the commercial vendor conformed to the regression model still warrants further investigation. It was likely that additional chemical heterogeneity such as the nonuniform distribution of the substitution along the cellulose backbone may play a role. It has been found that more heterogeneously substituted HPMC have impacted its solution properties such as cloud point[54] thus, may impact drug release from hydrophilic matrices.

Based on reviewing drug release mechanisms from hydrophilic matrices, the erosion rate of polymer was concluded to play a significant role in the drug release process. Therefore, a method was developed to measure the erosion rate of pure HPMC samples. The results confirmed that niacin release from the ER tablets correlated highly with the erosion rate of HPMC, regardless of the vendor (Fig. 36.6).

Based on the investigation, a specification was then set on this property of HPMC. With the improved control on HPMC, no further dissolution issue has surfaced. In addition to that, predicted and observed drug release has been consistent. The Niacin ER tablets drug release at 20 hours data from 311 commercial batches manufactured post the implementation of the change are presented in Fig. 36.7, and a process capability analysis plot is given in Fig. 36.8. The 95% lower confidence bound of the performance index (LB for $P_{pk} = 1.43$) indicates the additional control led to great

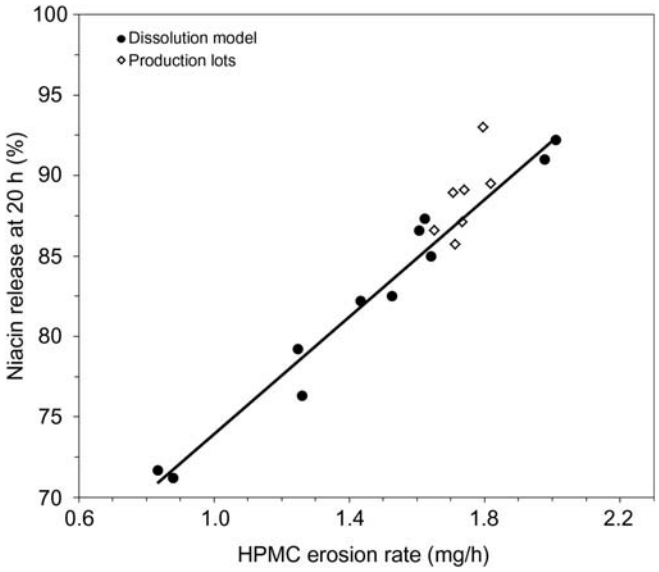


FIGURE 36.6 Correlation between niacin release at 20 hours and HPMC erosion rate.

improvement in meeting the quality requirement of drug release at 20 hours.

In the above example, the lot-to-lot and vendor-to-vendor variability of HPMC was not apparent from the typical compendia type of properties conventionally measured. In the end, a specification was developed based on measurement of an unconventional property of HPMC to control its variability. Searching beyond the compendia properties and understanding the dosage form design are necessary to understand better and control excipient variability. The case study also demonstrated that control chart can be an effective tool to detect the presence of special cause variation in the manufacturing process and to ascertain if the process has reached a state of statistical control. The control charts also serve as indicators of impending problems and allow operating personnel or process engineers to take proactive actions to prevent product quality failure from happening.

36.4.5 Illustrative Example-2: using PAT and RTRT to monitor and control traditional batch manufacturing process

CPV and PAT/RTRT can be applied independently. However, there are synergies between the two when they are applied together—for example, the ability to respond more quickly to data generated and monitored in real time (as opposed to waiting for traditional QC testing). Below is a schematic of conceptual example (Fig. 36.9) of using PAT/RTRT for monitoring and controlling a typical solid oral dosage manufacturing process using traditional batch manufacturing process.

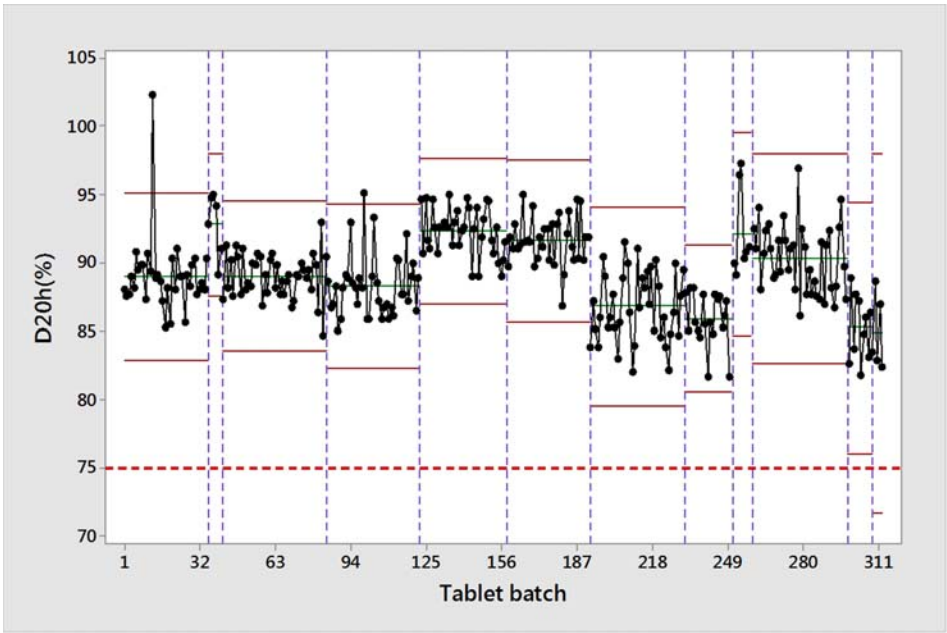


FIGURE 36.7 Run chart of niacin release at 20 hours, after the additional HPMC specification is implemented.

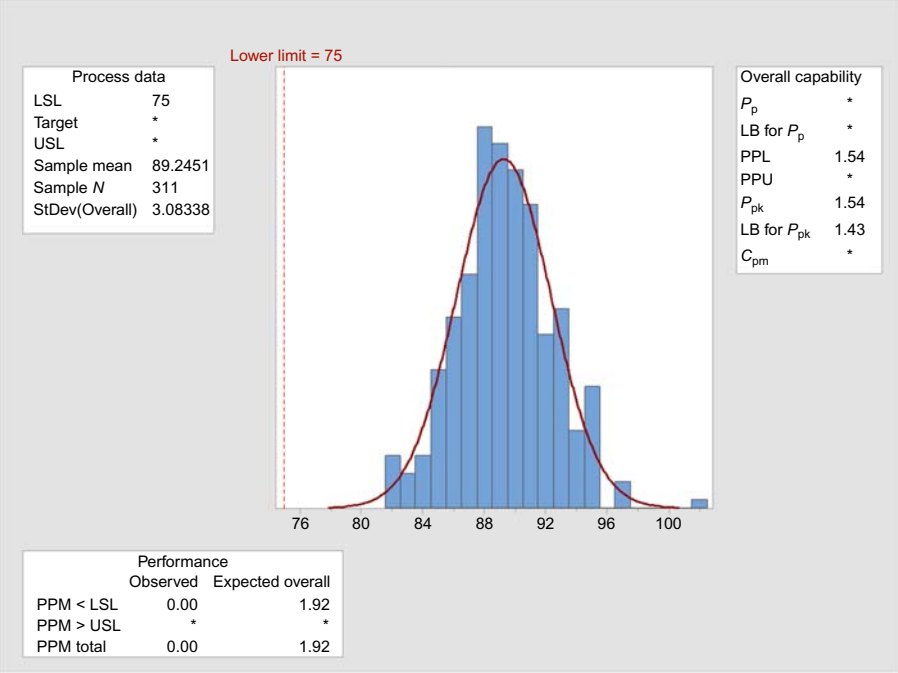


FIGURE 36.8 Process capability analysis of niacin release at 20 hours, after the additional HPMC specification was implemented.

The example is a typical solid oral dosage manufacturing process using traditional batch manufacturing process. The manufacturing steps include dispensing of raw materials, blending, milling, roller compaction, compression, and coating. Many on-line and in-line PAT tools are used including the use of in-line NIR for blend uniformity monitoring, in-line or on-line particle size monitoring by laser diffraction, and at-line monitoring of tablet active content and identity (ID) by NIR. RTRT is implemented for the following tests: assay, CU, ID, and dissolution. RTRT relies on the measurement from the PAT tools. In this example, tablet assay and CU for release are calculated based on at-line NIR measurement during compression, tablet identity for release relies on at-line ID measurement on tablet cores, multivariate mathematical model to predict dissolution for release is based on

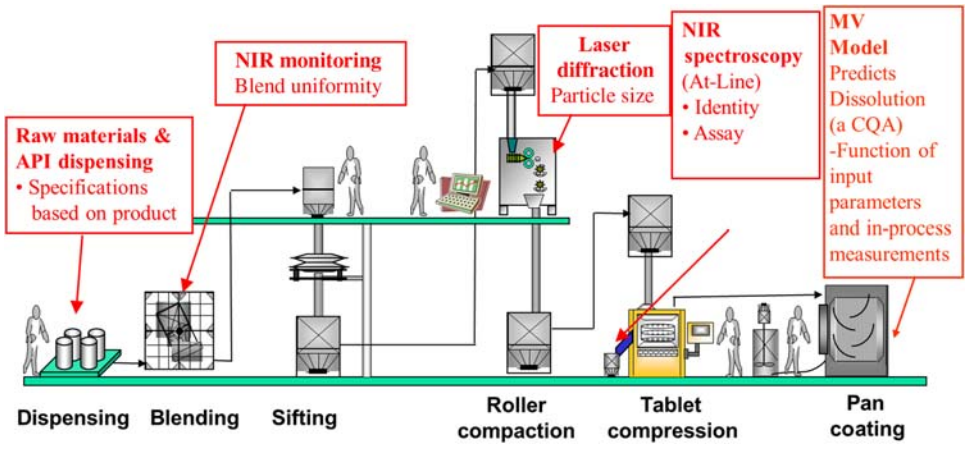


FIGURE 36.9 A schematic of conceptual example of PAT/RTRT implementation for a typical solid oral dosage manufacturing process using traditional batch manufacturing process.

TABLE 36.2 Highlights of Scientific Considerations for Implementing RTRT

Type of RTRT	Highlights of some specific questions to address during implementation
Real-time ID testing for incoming materials	• Is the ID test capable of distinguishing other raw materials available at the manufacturing location?
Real-time ID testing for finished products	• Is the ID testing capable of distinguishing between other products manufactured at the manufacturing location? • Does the product have any unique identifiers such as embossing that would interfere with the ID testing, depending upon orientation of the sample? • If the ID testing is performed on an intermediate instead of the finished product, what are the process steps after the point of ID testing that would disallow any potential human and/or system error?
Real-time measurement of core tablet content uniformity by NIR	• Is the sample size and sampling frequency for NIR measurement, representative of the batch? • Is there potential for segregation to occur in the feed hopper or feed frame after the point of measurement? • Since NIR measures active concentration in the tablet, is variation in weight of the tablet accounted for when calculating total active content in the tablet? Is the acceptance criteria proposed for content uniformity proportional to the sample size measured by NIR? • Is the RTRT method comparable to the traditional analytical method, for example, HPLC? • Is the calibration model associated with the NIR method adequately developed, validated, and maintained and is it applicable over the proposed range of operation at commercial scale?

Considerations for models serving as surrogates for release test	• Does the model include all the variables that have a potential to impact the attribute that is being modeled?
	• Is the model capable of detecting nonconforming batches?
	• Is the calibration model robust? For example:
	  ◦ Does it include possible variations in raw materials and processing conditions, that is expected to occur at commercial scale?
	  ◦ Has the model been validated using statistically approaches?

in-process measurements made during the production of material and process attributes that impact dissolution. Table 36.2 highlights some specific considerations (not exclusive) for different types of RTRT implementation for this case study in addition to the general considerations discussed in ICH Q8 (R2) and other ICH documents.[6,7]

If a model is used to support RTRT, as in the case where a model serves as a surrogate for release test, it is regarded as a high impact model, as per the criteria outlined in the "Models" section in the ICH QIWG *Points to Consider* document.[5] A model can be considered high the impact if the prediction from the model is a significant indicator of the quality of the product. Examples of such models can include multivariate models to predict dissolution for release, calibration models associated with NIR procedures that are used

for CU and assay release testing. For these cases, the ICH IWG "Points to Consider" document provides useful direction. Various examples of these models are also discussed in the paper on the role of models in pharmaceutical industry.[55–57]

36.4.6 Illustrative Example-3: using PAT and RTRT to monitor and control a continuous manufacturing process

Continuous manufacturing (CM) process refers to a manufacturing scheme where the material(s) and product are continuously charged into and discharged from the system, respectively, throughout the duration of the process.[58] CM is regarded as an innovation that has lots of potential to improve agility, flexibility, and robustness in the manufacture of pharmaceutical products. In addition, CM offers many other advantages such as smaller equipment footprint (and therefore CGMP processing area), shorter cycle times, inventory reduction, improved yields, reduced capital costs, and improved control over quality (http://www2.aaps.org/Arden/). More details regarding CM will be discussed in chapter "Emerging Technology for Modernizing Pharmaceutical Production: Continuous Manufacturing." This section focuses on the application of PAT/RTRT for monitoring and controlling CM process.

CM requires control systems that ensure continuous monitoring of quality of the intermediates and finished products, with the ability to detect and remove nonconforming batches. This leads to implementing many on-line/at-line PAT tools or models that can predict quality properties based on process data, that is, soft sensors, and information from which can then be leveraged to implement RTRT. Schematic below shows a conceptual example (Fig. 36.10) of implementing PAT/RTRT for a CM process via dry granulation for solid oral dosage form.

In this example, raw materials are added to a continuous blender via loss in weight feeders. The blended material then goes through a continuous dry granulation process such as roller compaction. Roller compacted ribbons are milled into granules which are then lubricated using a continuous blender where the lubricant is directly added to the blender using a loss in weight feeder. Lubricated granules are then compressed using a rotary tablet press and coated using a continuous film coating method. In this example, RTRT is implemented for the following tests: Assay, CU, ID, and Dissolution. RTRT for tablet assay and CU is calculated based on on-line NIR measurement during compression, tablet identity for release relies on on-line identity measurement on tablet cores, and a multivariate mathematical model is used to predict dissolution as a function of in-process measurements of active amount, tablet hardness, and granule particle size distribution. Additionally, for the continuous system, information from the in-line/on-line PAT tools can be used for process control. Data from PAT tools can be used either in feedback or feedforward mode to adjust the process parameters (as needed) to ensure that the process is operating as desired. To sum, real-time monitoring and release testing offers the advantage of continual assurance of quality during continuous manufacturing operation, however adoption of RTRT is not a requirement for CM implementation.

36.5 SUMMARY

Both the regulators and the pharmaceutical industry have been confronted with new and increasingly complex challenges, such as unprecedented drug shortages and recalls. By far, the most frequently cited reasons for drug shortages relate to manufacturing and product quality issues. Hence, a heightened focus is needed for better understanding the impact of commercial manufacturing

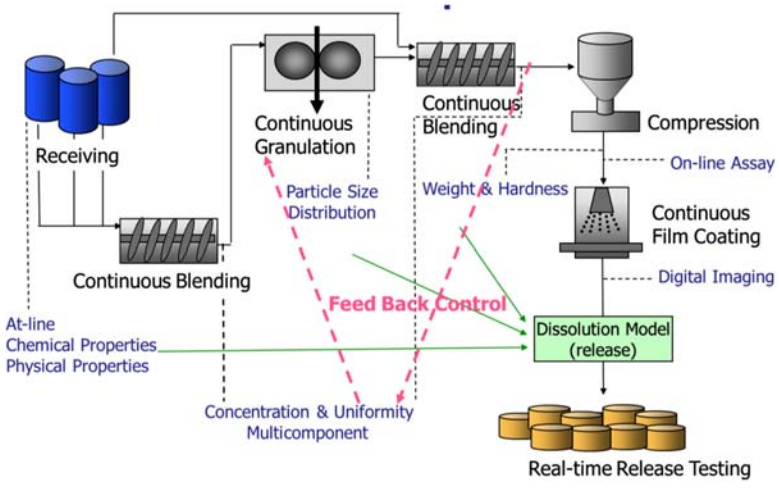


FIGURE 36.10 A schematic of conceptual example of PAT/RTRT implementation for a continuous manufacturing process.

on product quality. The topics covered in this chapter—lifecycle approach to process validation, routine commercial process monitoring, and continual improvement as well as PAT and RTRT applications—are part of the overall effort to improve pharmaceutical product quality and process performance.

References

1. Yu LX. From our perspective: patients deserve quality medications; 2015. http://www.fda.gov/Drugs/NewsEvents/ucm428298.htm (accessed June 25, 2015).
2. Yu LX, Woodcock J. FDA pharmaceutical quality oversight. *Int Journal of Pharm* 2015;**491**:2–7.
3. U.S. FDA Guidance for Industry on Process Validation: General Principles and Practices, 2011.
4. European Medicines Agency. Guideline on process validation for finished products—information and data to be provided in regulatory submissions; 2014. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/02/WC500162136.pdf (accessed June 25, 2015).
5. <http://www.urbandictionary.com/define.php?term=Manufacturability>.
6. ICH Q8 (R2). Pharmaceutical development; 2009.
7. ICH. Quality implementation working group points to consider (R2): ICH-Endorsed Guide for ICH Q8/Q9/Q10 Implementation; 2011.
8. ICH Q9. Quality risk management; 2005.
9. ICH Q10. Pharmaceutical quality system; 2008.
10. Yu LX, Amidon G, Khan MA, Hoag SW, Polli J, Raju GK, et al. Understanding Pharmaceutical Quality by Design. *AAPS J* 2014;**16**:771–83.
11. U.S. Food and Drug Administration. Guidance for industry: PAT—a framework for innovative pharmaceutical development, manufacturing, and quality assurance; 2004.
12. Simon LL, Pataki H, Marosi G, Meemken F, Hungerbuhler K, Baiker A, et al. Assessment of recent process analytical technology (PAT) trends: a multiauthor review. *Org Proc Res Dev* 2015;**19**:3–62.
13. Yu LX. Pharmaceutical quality by design: product and process development, understanding, and control. *Pharm Res* 2008;**25**:781–91.
14. Lionberger RA, Lee SL, Lee L, Raw A, Yu LX. Quality by design: concepts for ANDAs. *AAPS J* 2008;**10**:268–76.
15. Raw AS, Lionberger R, Yu LX. Pharmaceutical equivalence by design for generic drugs: modified-release products. *Pharm Res* 2011;**28**:1445–53.
16. U. S. Food and Drug Administration. Quality by design for ANDAs: an example for immediate-release dosage forms; 2012. http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/AbbreviatedNewDrugApplicationANDAGenerics/UCM304305.pdf. (accessed June 25, 2015).
17. U. S. Food and Drug Administration. Quality by design for ANDAs: an example for modified-release dosage forms; 2011. http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/AbbreviatedNewDrugApplicationANDAGenerics/UCM286595.pdf. (accessed June 25, 2015).
18. Product Quality Lifecycle Implementation (PQLI®) from Concept to Continual Improvement Part 1- Product Realization using Quality by Design (QbD): Concepts and Principles. International Society for Pharmaceutical Engineering (ISPE), November, 2011.
19. Product Quality Lifecycle Implementation (PQLI®) from Concept to Continual Improvement Part 2- Product Realization using Quality by Design (QbD): Illustrative Example. International Society for Pharmaceutical Engineering (ISPE), November, 2011.
20. Product Quality Lifecycle Implementation (PQLI®). Product development and realization case study A-mab; 2009. http://www.ispe.org/pqli/a-mab-case-study-version-2.1.pdf. (accessed June 25, 2015).
21. Pharmaceutical Development Case Study: ACE Tablets. https://www.ispe.org/pqli/case-study-ace-tablets.pdf. (accessed June 25, 2015).
22. 2.3 Quality Overall Summary-Sakura Tablet. 2008. www.nihs.go.jp/drug/PhForum/FileCEnglish.pdf. (accessed June 25, 2015).
23. Technical Report No. 60: Process Validation: A lifecycle approach; Parenteral Drug Association; 2013.
24. ASTM E2500-07. Standard Guide for specification, design, and verification of pharmaceutical and biopharmaceutical manufacturing systems and equipment; American Society for Testing and Materials; 2007.
25. ISPE Baseline Guide: Commissioning and Qualification, Vol. 5; International Society for Pharmaceutical Engineers; 2001.
26. ISPE Baseline Guide: Risk-based manufacture of pharmaceutical products (Risk-Mapp); International Society for Pharmaceutical Engineers; 2010.
27. Wiles F. Risk-based methodology for validation of pharmaceutical batch processes. *PDA J Pharm Sci Technol* 2013;**67**:387–98.
28. Yang H. How many batches are needed for process validation under the New FDA Guidance? *PDA J Pharm Sci Technol* 2013;**67**(1):53–62.
29. ISPE Discussion Paper—Topic 1-Stage 2 Process Validation: Determining and Justifying the Number of Process Performance Qualification Batches. http://www.ispe.org/discussion-papers/stage-2-process-validation.pdf. (accessed June 25, 2015).
30. Makkonen L. Extreme value analysis and order statistics, bringing closure to the plotting position controversy. *Commun Statis Theory Methods* 2008;**37**:460–7.
31. Kakad S., Wale K., Balsane M., Salunkhe K.S., Chaudhari S.R., Process validation of parenteral formulation. World J Pharm Res, 3(Suppl. 5): 241-253.
32. Yu LX, Peng DY, Lionberger R, Viehmann A, Iyer K. Using process capability to ensure product quality. *Pharm Eng* 2015;**35**(2):35–43.
33. ISO 16269-6:2014, Statistical Interpretation of Data, Part 6 - Determination of Statistical Tolerance Intervals (Geneva, Switzerland, Jan. 23, 2014).
34. ISPE Discussion Paper- Topic 4- Evaluation of Impact of Statistical Tools on Process Performance Qualification (PPQ) Outcomes. http://www.ispe.org/discussion-papers/statistical-tools-ppq-outcomes.pdf. (accessed June 25, 2015).
35. Mitchell M. Determining criticality–process parameters and quality attributes part III: process control strategies—criticality throughout the lifecycle. BioPharm International, 27, 3.
36. ASTM E2281: standard practice for process and measurement capability indices.
37. ISO 21747: Statistical methods—process performance and capability statistics for measured quality characteristics.
38. ISO 11462: Guidelines for implementation of statistical process control (SPC).
39. Montgomery DC. *Introduction to statistical quality control*. 7th ed New York, NY: Wiley; 2012.
40. Bothe DR. *Measuring process capability*. Cedarburg, WI: Landmark Publishing Inc.; 2001.

41. Botelho TS, Tavares VF, Curtivo CPD, Sarolli SRB, Fernandes MA, donaduzzi CM, et al. A statistical approach to evaluating the manufacture of Furosemide tablets. *Pharm Tech* 2011;**35**(3):112–21.
42. Chopra V, Bairagi M, Trivedi P, Nagar M. A case study: application of statistical process control tool for determining process capability and sigma level. *PDA J Pharm Sci and Tech* 2012;**66**(2):98–115.
43. Chabukswar AR, Jagdale SC, Kuchekar BS, Joshi VD, Deshmukh GR, Kothawade HS, et al. Six sigma: process of understanding the control and capability of Ranitidine hydrochloride tablet. *J Young Pharm* 2011;**3**(1):15–25.
44. Peng DY, Lionberger R, Viehmann A, Iyer K, Yu LX. Using control chart to evaluate pharmaceutical manufacturing process variability. *Pharm Eng* 2015;**35**(3):59–74.
45. Statistical Control Handbook, 1st Ed., 1956, by Western Electric Company.
46. Barbosa de Alencar JR, Lopes CE, Bezerra de Souza Jr. M. Multivariate statistical process control: monitoring application to captopril tablet production. *Revista Brasileira de Farmacia* 2007;**88**:200–5.
47. Kourti T. Quality by design in the pharmaceutical industry: the role of multivariate analysis. *Am Pharm Rev* 2009;**12**:118–23.
48. Laursen K, Rasmussen MA, Bro R. Comprehensive control charting applied to chromatography. *Chem Intel Lab Sys* 2011;**107**:215–25.
49. Kona R, Qu H, Mattes R, Jancsik B, Fahmy RM, Hoag SW. Application of in-line near infrared spectroscopy and multivariate batch modeling for process monitoring in fluid bed granulation. *Int J Pharm* 2013;**452**:63–72.
50. Peng DY, Lostritto R, Bika D, Geoffroy J, Shepard T, Eden B, et al. Symposium summary report: the use of process capability to ensure pharmaceutical product quality. *Pharm Eng* 2014;**34**(5):10–23.
51. Peng DY, Zilian A, Norton J, VanTrieste M, Orloff J, Stojanovski P, et al. Symposium summary report: using process capability to enhance pharmaceutical product quality. *Pharm Eng* 2015;**35**(5):126–31.
52. De Feo JA, Barnard W. *Juran Institute's six sigma breakthrough and beyond—quality performance breakthrough methods*. Noida, UP: Tata McGraw-Hill Publishing Company Limited; 2005.
53. Zhou D, Law D, Reynolds J, Davis L, Smith C, Torres JL, et al. Understanding and managing the impact of HPMC variability on drug release from controlled release formulations. *J Pharm Sci* 2014;**103**:1664–72.
54. Viridén A, Wittgren B, Larsson A. Investigation of critical polymer properties for polymer release and swelling of HPMC matrix tablets. *Eur J Pharm Sci*. 2009;**36**(2-3):297–309.
55. Kourti T. Process analytical technology beyond real-time analyzers: the role of multivariate analysis. *Crit Rev Anal Chem* 2006;**36**:257–78.
56. Kourti T. Chapter 44: Pharmaceutical manufacturing: the role of multivariate analysis in design space, control strategy, process understanding, troubleshooting, and optimization. In: am Ende David J, editor. *Chemical Engineering in the Pharmaceutical Industry*. Hoboken, New Jersey: John Wiley & Sons, Inc; 2011. p. 853–78.
57. Kourti T, Lepore J, Liesum L, Nasr M, Chatterjee S, Moore CMV, et al. Scientific and regulatory considerations for implementing mathematical models in the quality by design (QbD) framework. *Pharm Eng* 2014;**34**(6):1–21 (Online Exclusives) (accessed on June 25, 2015) (http://www.ispe.org/pharmaceutical_engineering/november2014)
58. Lee S, O'Connor T, Yang X, Cruz C, Chatterjee S, Madurawe R, Moore CMV, Yu L, Woodcock J. Modernizing pharmaceutical manufacturing: from batch to continuous production. *J Pharm Innov* 2015;**10**(3):191–9.

CHAPTER

37

Emerging Technology for Modernizing Pharmaceutical Production: Continuous Manufacturing

T. O'Connor and S. Lee

U.S. Food and Drug Administration, Washington, DC, United States

37.1 INTRODUCTION

Over the past decade, there has been progress in the adoption of systematic science and risk approaches for pharmaceutical development, which emphasize product and process understanding and control strategy. Nevertheless, available product quality data show that pharmaceutical manufacturing and its regulation, in general, are being confronted with new and increasingly complex challenges, such as the unacceptably high occurrence of product recalls and drug shortages,[1,2] globalization of drug product manufacturing that results in complex supply chains that span several countries, and lack of manufacturing flexibility for optimization and continuous improvement, resulting in an increasing number of postapproval supplements for manufacturing changes submitted to the US Food and Drug Administration (FDA) for review.

The adoption of emerging technology is one of the key mechanisms for addressing these challenges. Modernizing manufacturing technology may lead to a more robust manufacturing process with fewer interruptions in production, fewer product failures (before or after distribution), and greater assurance that the drug products manufactured in any given period of time will provide the expected clinical performance. Continuous pharmaceutical manufacturing is an emerging technology that offers potential efficiency, agility, flexibility, and quality advantages over traditional processing methods.[3] A number of technologies are available for continuous drug product manufacturing. There are inherent implementation challenges, but the extent of the challenges will depend upon the specific product and the manufacturing route selected. This chapter describes some of the current challenges facing pharmaceutical manufacturing. Next, it discusses how the adoption of emerging technology can enable the addressing these challenges, using continuous manufacturing as an example. Then technologies that enable continuous drug product manufacturing, the so-called hybrid continuous models for powder processing, are described, and finally implementation challenges are highlighted.

37.2 CHALLENGES FOR PHARMACEUTICAL MANUFACTURING

Pharmaceutical manufacturing processes are generally less efficient and understood compared with those in other chemical process industries.[4] This has resulted in product recalls and failures due to inherent defects in product and process design. The number of Class I recalls, which denotes a situation where there is a reasonable probability that exposure to a defective product will cause serious adverse health consequences or death, have increased sevenfold over the past decade.[1] Examination of the data further indicates that these recalls are typically due to failures in the implementation of manufacturing process scale-up and routine production. For example, in 2014, the chief causes of pharmaceutical recalls were the presence of particulate matter and failure to meet the product quality specifications.[5]

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00037-6

The lack of agility, flexibility, and robustness in the pharmaceutical manufacturing sector poses a potential public health threat because manufacturing failures may lead to poor product quality and drug shortages. Many drug shortages have been caused by the use of outdated equipment, reliance on aging facilities operating at the maximum production capacity, and lack of effective quality management systems. A 2012 FDA analysis of data collected form manufacturers determined that 66% of production disruptions leading to shortages resulted from either stoppages to address product-specific quality failures or general efforts to remediate a problematic manufacturing facility.[2]

Globalization continues to pose challenges for pharmaceutical manufacturing. Due to economic factors, supply chains for many active pharmaceutical ingredients (APIs) and final drug products span several countries and thereby contain multiple supply vulnerabilities. Under current manufacturing practices, intermediates may not be immediately processed. Instead, they are stored in containers and shipped around the world to the next manufacturing facility. This is a significant risk for APIs or intermediates that can degrade over time or are sensitive to environmental conditions, directly impacting the overall drug product quality. This trend is evident in the product recall data: 26% of drug recalls reported during 2014 to the FDA were international in nature, impacting more than one country, compared with just 4% in the beginning of 2013.[5]

For existing processes, the current regulatory framework may not help the implementation of manufacturing innovations generated from knowledge obtained during production. This effect is evidenced by the increasing number of postapproval supplements received for review over the past decade, in part owing to our current practice of "locking in" an applicant's manufacturing process before it is fully optimized. A burdensome regulatory framework causes manufacturers to submit supplements as they strive for process optimization. The large number of postapproval supplements also creates a drain on FDA resources that could otherwise be dedicated to other public health priorities. Due to the fractured global regulatory environment, the timeline between filing the first regulatory submission and final global approval for the implementation of a manufacturing improvement can be several years. Over this time period, the manufacturer must utilize different manufacturing processes in the various regions to supply patients with products. This creates supply chain and logistical challenges and vulnerabilities and slows the implementation of innovative technologies, which could improve product quality and availability.

37.3 THE ADOPTION OF EMERGING TECHNOLOGY TO ADDRESS PHARMACEUTICAL MANUFACTURING CHALLENGES

The adoption of emerging manufacturing technologies is one way to address the cited challenges. An emerging technology can be viewed as a technology that has the potential to modernize the pharmaceutical manufacturing body of knowledge to support more robust, predictable, or cost-effective processes. Examples of such elements include an innovative or novel: (1) product manufacturing technology (eg, dosage forms and delivery systems); (2) manufacturing process (eg, design, scale-up, and automation); and/or (3) testing and monitoring technology. As explained next, the adoption of emerging technologies can help drive the sector toward the vision of a maximally efficient, agile, flexible pharmaceutical manufacturing sector that reliably produces high-quality drugs without extensive regulatory oversight.[6]

Continuous manufacturing is an emerging pharmaceutical technology.[3] This manufacturing technology is not a new concept per se, since it has been extensively utilized in the petroleum, chemical, and food industries (among many others) for several decades. However, there is currently limited experience with implementing this technology within the pharmaceutical industry. Batch manufacturing is traditionally utilized for the production of pharmaceutical products. In this type of process, materials from one step are processed over a period of time set by recipe, then usually tested off-line as per the in-process controls, and finally stored before they are sent to the next processing step. If the in-process material does not meet quality expectations, it may be discarded or, under certain circumstances, reprocessed before moving to the next process step.[7] In continuous manufacturing, materials produced during each process step are sent directly and continuously to the next step for further processing. Each processing step needs to reliably produce its intermediate output material or product with acceptable characteristics. Modifying the processing rate of particular unit operation(s) (eg, synthesis, crystallization, and blending drying) to achieve the desired quality may not be possible for continuous manufacturing because it may create disruption for downstream unit operations unless buffer capacity is designed into the system.

Continuous manufacturing, compared with batch manufacturing, thus generally involves a more highly integrated level of process design and operation so as to ensure continuous control of product quality attributes. There are different types of continuous manufacturing implementation models. Fig. 37.1

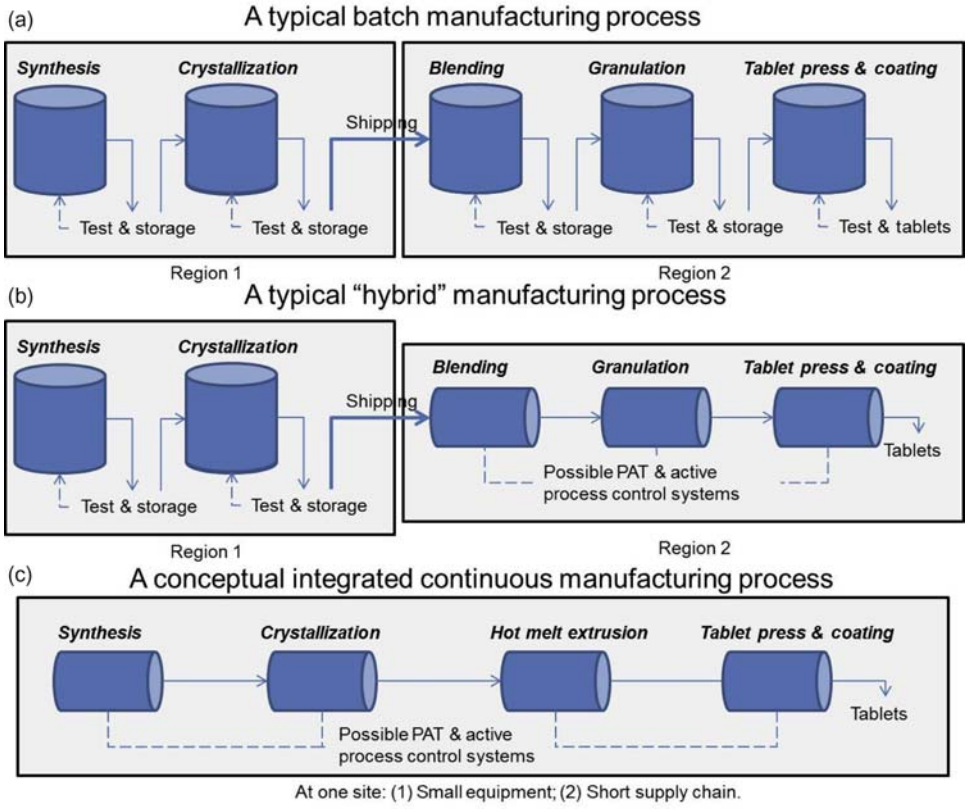


FIGURE 37.1 Conceptual examples of continuous manufacturing implementation. (a) An example of batch manufacturing process where the API and final dosage form are produced at different facilities and off-line testing is used to verify product quality between processing steps, (b) An example of a hybrid approach where the API is produced using batch technology, but the drug product starting with the API is manufactured continuously and supervised by an active control system that ensures quality, (c) An example of end-to-end continuous manufacturing process from raw materials to final dosage form. To take advantage of the unique aspects of continuous, end-to-end processes may lead to the utilization of emerging processing technologies, such as spray drying, hot melt extrusion, and printing, among others.

depicts two such implementation models (ie, hybrid and end-to-end) for continuous pharmaceutical manufacturing. Fig. 37.1b shows the drug product hybrid model where the API is produced separately using batch technology, and the drug product, starting with the API, is manufactured via an integrated continuous process. The hybrid model can also be applied to drug substance manufacturing where the API starting from raw materials is produced via an integrated continuous process, and the drug product is manufactured separately. Fig. 37.1c shows an example of end-to-end models where the continuous process starts with raw materials and produces the final dosage form. To take advantage of the unique aspects of continuous manufacturing, end-to-end processes may lead to the development and utilization of emerging process technologies. The future vision for continuous manufacturing for both models is one in which: (1) individual continuous unit operations are connected to form an integrated manufacturing process, (2) process analytical technology (PAT) systems are utilized to provide real-time data for process monitoring and control, and (3) engineering process-control systems are implemented to mitigate the impact of raw material and process variability on the quality of finished products. Compared with traditional batch manufacturing, continuous manufacturing provides several potential opportunities to improve control of product quality and thereby increase the efficiency, flexibility, and agility of manufacturing, as highlighted next.

Continuous unit operations are generally more efficient than their batch counterparts and offer much higher throughput per unit volume and per unit time, thereby often greatly reducing the size of the processing equipment. The manufacturing footprint is further reduced because the material flows from one processing step to the next, and does not need isolated suites or dedicated modules. For this reason, a substantial reduction in both capital and operating expenses for a continuous process can be achieved.[8] Furthermore, continuous processing can decrease the amount of potentially expensive API required for process development studies, greatly

reducing the material cost of process development and optimization efforts.

Continuous manufacturing may also increase process efficiency by facilitating the streamlining of the manufacturing process through the removal of corrective or work-up unit operations. One potential example of process intensification (collapsing of a multistep process into a smaller number of steps or unit operations) is reaction that telescopes drug substance manufacturing.[9] Traditional batch reactions normally include several reaction steps with isolation and purification work-up operations following each reaction step. The postreaction work-up steps can be highly time consuming and generate large volumes of waste. To avoid this lengthy process, telescoping could be an ideal alternative and well-suited for continuous flow chemistry.[10] Other examples include: (1) the potential to eliminate downstream corrective drug product unit operations, such as sizing steps (eg, granulation, milling), due to better control of the particle size distribution during crystallization, and (2) reduction in the segregation risk from continuous blending.[11,12] Continuous manufacturing may also facilitate the adoption of emerging processing technologies which are well-suited for continuous operations.[13]

Continuous manufacturing also has the potential to improve the agility and the flexibility of the pharmaceutical manufacturing sector. The pharmaceutical industry currently has a limited ability to rapidly increase production in the face of drug shortages or other emergencies, such as pandemics. Bringing up a new facility or manufacturing line in response to such emergencies may take up to several months or even years. Continuous manufacturing can potentially permit increasing production volume without the current obstacles related to scale-up, providing more response capacity. Scale-up options, such as operating the process for longer periods of time, utilizing parallel processing lines, or increasing the flow rate through the process, can be built into the process design and verification. Modularized skid-mounted continuous plants also offer the possibility of rapidly bringing additional manufacturing capacity on-line.[14] In addition, due to the small volume of materials needed to run continuous manufacturing systems, it may be possible to design and optimize a continuous system on commercial scale equipment, thus eliminating scale-up.[15,16] Eliminating scale-up bottlenecks in the path to market may facilitate rapid clinical development of breakthrough drugs.

Continuous manufacturing provides advantages to address complex supply chain issues due to globalization. Continuous manufacturing allows the production at various scales with a given process, which may facilitate regional or in-country manufacturing. Under a continuous operating mode, hold times between steps can be eliminated. This is a significant advantage for APIs or intermediates that can degrade over time or are sensitive to environmental conditions, directly improving the overall drug product quality. Furthermore, the small scale of continuous manufacturing can decrease the safety hazards associated with highly energetic or hazardous materials and potentially allow for more flexibility in the use of nonspecialized manufacturing facilities.[17]

Continuous manufacturing is strongly aligned with the FDA's support of the quality-by-design (QbD) paradigm for pharmaceutical development and offers opportunities to improve process robustness. Development of a robust process relies on utilizing product and process understanding to identify sources of variation to product quality and to design appropriate control strategies to address these risk areas. Continuous manufacturing provides an opportunity to utilize enhanced product and process understanding to adopt advanced manufacturing controls and produce uniformly high-quality products with reduced waste, resulting from the generation of out-of-specification material.[4]

37.4 TECHNOLOGIES FOR CONTINUOUS DRUG PRODUCT MANUFACTURING

The design of the manufacturing process for solid dosage forms varies from compound to compound because the properties of the API molecule play a significant role in the development of an appropriate formulation. The majority of solid oral drug products are produced using one of several manufacturing routes: direct compression, wet granulation, or dry granulation. Technologies are available to implement an integrated continuous process line for each of these design alternatives. Fig. 37.2 depicts the unit operations utilized for three continuous manufacturing routes.[18] New process technologies are also being developed or adapted to the pharmaceutical industry (eg, spray drying, hot melt extrusion (HME), casting, and 3D printing) to take advantage of continuous processing in order to combine the function of multiple steps into a single process or reduce solid handling.[13] These technologies can enable end-to-end manufacturing from raw materials to the final dosage form in a single, integrated continuous process.[16] In general, even if the drug substance and drug processes are not physically integrated, the processes should be designed as a single system, with a consistent design objective, so that downstream processes are not implemented to correct issues that could have been controlled upstream. Regardless of the manufacturing route, continuous manufacturing naturally lends itself to the adoption of

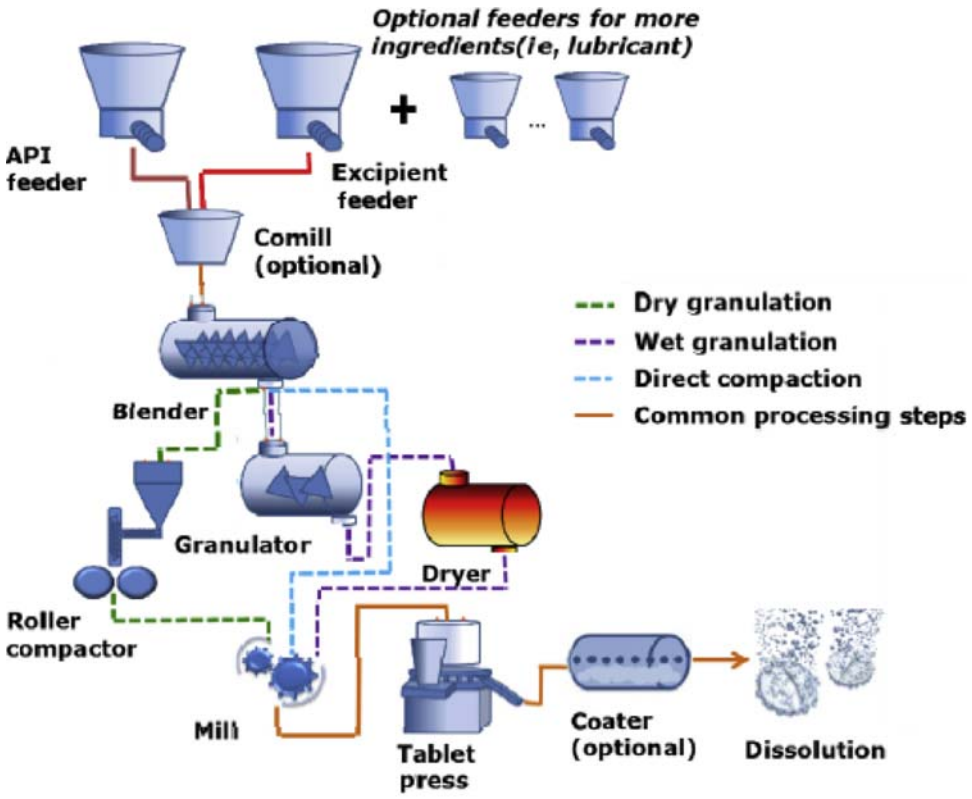


FIGURE 37.2 Continuous drug product manufacturing routes. Three manufacturing routes for the continuous production of pharmaceutical tablets are shown. Common processing steps for all three routes include feeding, blending, and tableting. The dry granulation route involves roller compaction followed by milling of the produced ribbons, while the wet granulation route involves wet granulation followed by drying and then by granule milling. In the case of direct compaction, there is no granulation step and therefore milling is not required prior to tableting.[18]

advanced approaches for real-time monitoring and process control due to the absence of isolated intermediates and the typically faster process dynamics for a continuous process that may necessitate more frequent measurements. These process monitoring and control systems are typically an essential component of the control strategy for a continuous process.

37.4.1 Feeding

The objective of the feeding operation is to supply the correct ratio of raw materials specified by the drug product formulation. In a continuous process, this requires feeding each material at a specified rate, such that the final product will have the proper composition. If the feed rate of one ingredient changes even for a brief period of time, the resulting variation in composition of the process stream can propagate downstream,[19,20] potentially leading to out-of-specification product units. The ability to feed powders consistently and continuously is typically accomplished through the use of loss-in-weight (LIW) feeders.

LIW feeders comprise a volumetric feeder mounted on top of a weighing platform that measures the mass of the feeder and its powder contents. The most common mechanism for feeding utilizes a hopper mounted on top of a positive-displacement screw element. When the screws are in motion, powder is fed from the material hopper into the process, and the total weight of the screw feeding system and the hopper decreases at a rate equivalent to the rate at which material is fed into the process. A gravimetric controller continuously monitors this LIW to adjust the screw speed so that the rate at which material enters the process remains on target. Alternatively, feeders can operate in volumetric mode, where the screw speed is kept at a constant value based on a predetermined calibration to achieve the desired feed-rate set point. Volumetric mode cannot adjust for those variations in powder density that typically occur when emptying and refilling the hopper[21]; however, when gravimetric control mode is not possible, such as during start-up and hopper refill, volumetric mode is necessary.

Even under gravimetric control, the granular nature of the materials places a limit on how accurately the

addition of material can be controlled. This is due to the fact that powders are composed of three phases: solid (ie, the particles), liquid (ie, moisture content attributed to water molecules on the surface or in between particles), and gas (ie, air between the particles); bulk powder behavior is dependent on how these three phases interact. Powders can be made to flow when energy is applied but then can hold shape when the energy is removed. This complex behavior leads to variations in the interaction between feeding screw and material, both within the screw and at the exit of the feeder, thus leading to variations in the feeding rate. The amplitude and frequency of feeding variations depends on the material properties of the powder and the feeder design and configuration.[19]

A general methodology for design of the feeding operation has been outlined based on the following steps: select potential feeders and tooling (eg, screw type and exit screen) based on the desired volumetric throughput; screen the potential tooling options for compatibility issues (eg, material adhering within the flights of the screws and clogging of screens with cohesive material); and test compatibility and monitor feed rate. During feeding trials, selection of the proper sampling interval is important because a sampling interval that is too long relative to the subsequent processing will not adequately detect transient fluctuations or disturbances.[19,20] At high feeding capacities, the discrete nature of the powder stream becomes less significant and feeding accuracy is greatly improved. The main challenge lies in feeding solid materials accurately at low-flow rates, where the variations can become large compared with the rate at which the material is being fed. Minor components (eg, lubricants and disintegrants) are thus more susceptible to feeding limitations.[13] This limitation is also generally compounded by the powder properties of these materials (eg, low density material, such as SiO_2, and material with poor-flow properties). However, precision feeders are becoming available, which can mitigate the challenges posed by low-flow rates.

37.4.2 Blending

The purpose of blending in downstream pharmaceutical applications is to reduce composition variability and thus minimize spatial and temporal composition variations in the material exiting the mixer. Convective mixers are the most common type of continuous blender, comprising a horizontal or inclined tube with a bladed impeller running down its central axis, whereby the rotating blades cause convective mixing. Design parameters for this type of blender are tube length and diameter, and blade size, shape, and configuration.[18] Material and operating parameters that can impact performance are impeller rate, flow rate, bulk density, and cohesion.[22] A well-designed continuous mixer should be able to produce evenly distributed (not segregated) blends with good control of composition over time. Achieving this objective can be described in terms of radial and axial mixing.[23] The degree of radial mixing varies as a function of distance traveled down the length of the mixer. Thus, a snapshot of a radial cross-section of the blender tube near the entrance at steady state would show two unblended powders. If the blender is properly designed and operated, a snapshot of the radial cross-section near the exit of the blender would show two powders blended together (Fig. 37.3).[12] Radial mixing is a steady state process, and can largely be considered time-invariant as opposed to a batch-mixing process, where the homogeneity of the powder blend varies as a function of time.

Axial mixing (ie, back-mixing) refers to mixing along the length of the tubular mixer. As discussed in the feeding section, the feed rate of the drug product components will vary around the desired flow rates. If a continuous feeder simply radially blended the incoming materials, then any disturbance from feeding would pass right through the blender and end up as variation in the final drug product.[23] Axial mixing can reduce the impact of component feed rate fluctuations at the exit of the mixer because it categorizes blending between powder segments that enter the mixer at different times. In a continuous process, it is critical to

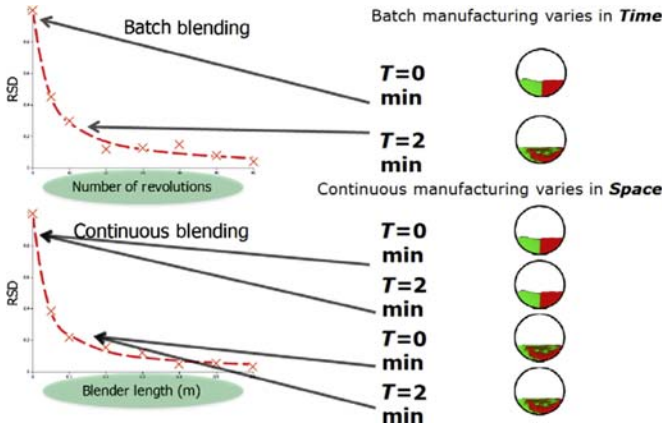


FIGURE 37.3 Comparison between batch and continuous blending. Difference in perspective for batch and continuous blending. The charts depict the relative standard deviation (RSD) for the content uniformity of a powder blend as a function of time for batch mixer and as a function of blender length for a continuous mixer. For the continuous mixer, the content uniformity does not vary as a function of time as shown by the callouts. Thus, the RSD for samples taken from the sample location in a continuous mixer should not change with time (ie, the samples taken at $T = 0$ min and $T = 2$ min are equivalent).[12]

understand that the feeding and blending systems must be designed to work together as a system. The amplitude and frequency of variations in the component feed rates must be fully understood because this leads directly to the degree of axial mixing that is required in the blending step. Impeller speed and cohesion have shown a significant interacting effect on axial mixing with cohesiveness, increasing axial mixing at high impeller rates, but having a negligible effect at lower impeller rates.[22] As a consequence of axial mixing, the particles in the powder blend will experience a distribution of processing times. This residence time distribution (RTD), a probability distribution that describes the amount of time a solid or fluid element is likely to remain in a particular unit or process, can be used to characterize a continuous mixer (eg, filterability of incoming disturbances), and the RTD will depend upon the design and operating parameters of the mixer as well as the physiochemical properties of the powder blend.[24] Process modeling, including RTD modeling, is expected to play a major role in determining the balance between axial and radial mixing that results in the desired product quality attributes, given the specific material feeding characteristics, for the drug product.[23]

Another distinguishing characteristic between a batch- and continuous-blending process is that the risk of segregation (ie, demixing) is lower in a continuous process. Segregation occurs when particles separate due to differences in their size, shape, or density. In a batch manufacturing process, the entire "batch" is present in the mixer for the entire processing time, while only a small portion of the batch is present in a continuous mixer at given moment in time. Thus, the radial length scale over which mixing occurs in the continuous process is much less, increasing mixing efficiency. Whereas in continuous processing, the steady-state material balance describes that the flux of material entering the system must equal the output flux, independent of particle size and density, otherwise there would be an accumulation of either smaller or larger particles in the blender. There is the possibility, however, that the RTD of different size particles could be different during dynamic phases of the process, such as start-up and shutdown.

37.4.3 Granulation

37.4.3.1 Wet granulation

Continuous wet granulation can be performed using a variety of techniques and equipment, with the most common including fluid-bed granulators and extrusion. The most widespread continuous fluid-bed granulators are horizontal moving-bed granulators/driers, which are mainly used in the chemical, dairy, and food industries. Although the concept of continuous moving-bed granulators can be used within the pharmaceutical industry, there are hardly any applications of them for the production of solid dosage forms, mainly because they are ideally suited for high-volume products. Another category of fluid-bed granulators is the rounded bed design, where the material is confined to an enclosed space (similar to batch fluid-bed processors), but able to continuously discharge agglomerated material through an outlet at the bottom of the screen (round-shaped). This allows the equipment to function over a greater range of production rates. A counter-current air flow through the pipe positioned at the center of the bottom plate ensures that only agglomerated particles are discharged. Only larger particles are able to overcome the air velocity, whereas the counter-current airflow carries undersized particles back into the processing chamber until they are sufficiently agglomerated. This allows the process to manufacture an essentially dust-free product, the velocity of the counter airflow determining the particle size of the end product.[25]

Extrusion is the most studied continuous granulation process for pharmaceutical application. This is due, in part, to the fact that extrusion is readily scaled-up and therefore lends itself to the pharmaceutical development process where the amount of product needed increases as it moves through the clinical development stages. Pharmaceutical granulation is often carried out in twin-screw extruders with co-rotating screws, an example of which is depicted in Fig. 37.4. In an extrusion granulation process, liquid binder is added continuously into the barrel and the moving powder is wetted and forms an extrudate. This extrudate is then broken into smaller pieces due to the shearing force that exists during the process. The material is then transported by conveying elements into mixing zones where it is mixed, compacted, and elongated by the kneading disks of the mixing zones. Mixing and agglomeration occur as the material is conveyed along the length of the barrel. A drying step is also required. In these processes, the main design parameters include the length of the barrel, the number of screws (one or two), the geometry of the screws, and the location of a point or points along the barrel where binding solution is added. Extruders used for granulation are operated in an open manner without a die at the end of the barrel. Granulation involves different rate processes, such as nucleation, coalescence, and consolidation (growth) and breakage. As opposed to batch granulation processes, where granulation rate processes may occur simultaneously, a dominant

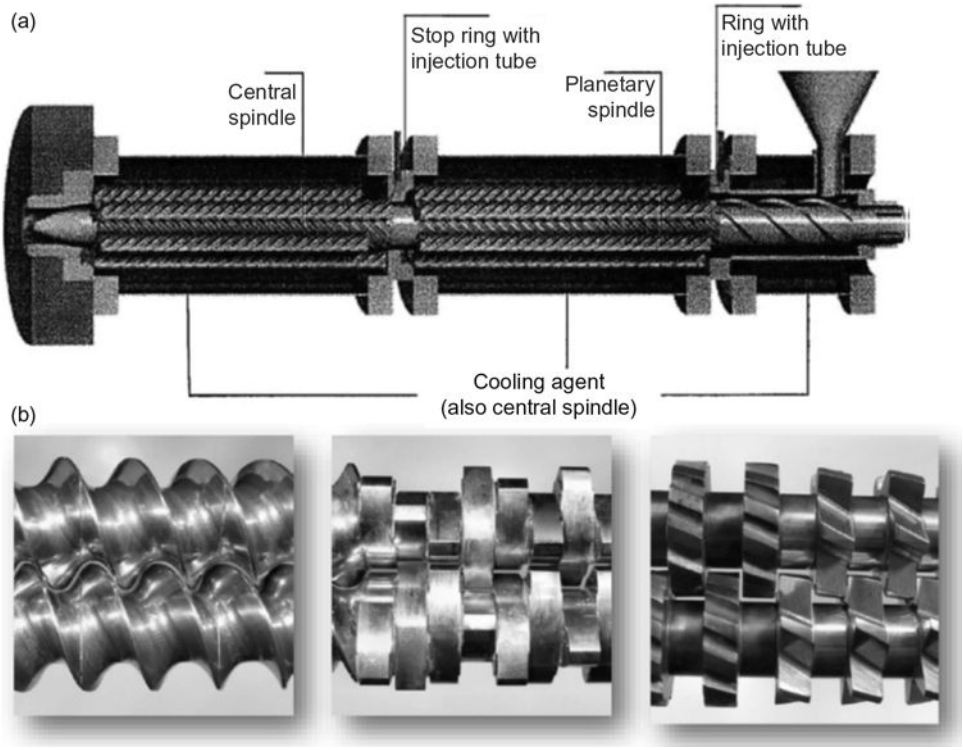


FIGURE 37.4 Twin-screw, high-shear wet granulator. (a) Twin-screw extruder, (b) Examples of co-rotating screws.[25]

granulation mechanism typically exists in each of the compartments along the length of the extruder.[26] Separating the granulation rate processes along the extruder should lead to better control of granule properties such as size and density, however, it is generally observed that extrusion tends to produce granules with a multimodal size distribution and greater porosity, usually requiring postprocessing size reduction. The broader size distributions observed are primarily due to relatively large droplet size of granulating liquid, liquid addition method, and poor liquid distribution compounded by the relatively short residence time in the granulator. Improvements in extruder design may be able to mitigate these challenges.[27]

It is important to understand the role of different screw elements (such as conveying and kneading elements) in order to describe the various rate processes in twin-screw granulator. Conveying screw elements possess axial displacement properties that are responsible for transporting the feed powder and liquid materials into the kneading section. The kneading elements provide a shear-intensive zone that allows granule formation by intimate mixing of liquid and powder; however, the granules obtained are typically characterized by a broad bimodal size distribution, with the presence of lumps and ungranulated fines.[26,28] Studies suggest that the compacted materials are also chopped by the conveying screws, which leads to the irregular and porous granules observed, and poor liquid distribution from dripping the binder liquid into the conveying section results in the observed bimodal broad granule size distribution.[28,29] Distributive mixing elements (DME) have also been studied for their use in twin-screw extrusion granulation. This type of screw element produces distributive mixing through cutting and recombination to allow collisions between in-flowing granular materials. Depending on their configuration relative to the incoming material, the DME elements produced granules with a narrower size distribution and improved liquid distribution.[30] This suggests the ability to tailor particle morphology and the potential to eliminate downstream milling operations.

Operating parameters of interest include screw speed, granulation temperature, and liquid-to-solid ratio.[31] Powder properties and binder properties, including viscosity and surface tension, have the potential to influence the process as well.[32,33] Tu et al.[34] have experimentally studied granule properties as a function of operating conditions, including screw speed and liquid-to-solid ratio. The authors developed regime maps for the process, describing the extruder geometry in terms of granulation (ie, the nucleation and growth of granules), extrusion (ie, material shaped into a continuous form, such as fiber or film), and breakage (ie, breaking of agglomerates into smaller

particles) regimes. The resulting regime maps suggest that desired granulation performance (narrow granule size distribution and large mean granule size) can be achieved with a higher liquid-to-solid ratio and a higher screw speed. However, increasing the liquid-to-solid ratio must be done with caution to avoid rapid growth and densification of granules beyond the processing ability of downstream units. The uniformity of size distribution can be improved by reprocessing and/or more aggressive mixing. Extrudates produced above a certain liquid-to-solid ratio tend to have a paste-like consistency.[34]

Granule drying is an inherent part of the wet granulation process. Drying is typically carried out utilizing a fluidized bed,[18] although other approaches, such as squeeze drying, belt drying, and drying through a screw, may be developed.[13] The fluid-bed drying process can be controlled through varying the air temperature, air velocity, product feeding rate, or the height of the powder in the bed, which can be adjusted by a weir installed at the outlet. The result of a fluid-bed drying process strongly depends on the RTD within the bed. The average residence time in the dryer significantly affects the propensity for undesired secondary agglomeration to occur during the drying process.[35] Thus, the key challenge in this operation is to achieve sufficient drying in a reasonable amount of residence time.

37.4.3.2 *Dry granulation*

Continuous granulation can also be accomplished via dry granulation through roller compaction. In the roller compaction process, powder is fed into a set of counter-rotating rolls, which compress and convey the powder forward with the motion of the rolls. Where the rolls are closest together, a compact ribbon is formed, which is continuously fed to a mill that transforms the ribbon into granules. Roller compaction can be a viable processing approach for powders with good flowability and compressibility, but may not perform well for powders with low bulk density, small particle size, or that are very cohesive. The development of a robust process can thus be challenging due to the strong dependency on raw material properties, however, because roller compaction results in the removal of the drying step, capital and operating costs are reduced and it can be used for moisture-sensitive APIs.[18]

Roller compaction equipment for a continuous process does not differ much from current batch production methods. Roller compactors vary mostly in terms of configuration. The orientation of rolls can be arranged either horizontally (with ribbons coming out parallel to the floor), vertically (with ribbons coming out perpendicular to the floor), or at some intermediate angle. In addition, the gap between the rolls can be fixed throughout the process or adjusted to account for variation in the incoming material. The roll diameter and width vary from one piece of equipment to another. The surface of the rolls can also vary from smooth to rough with grooves to pocketed designs. Finally, the powder can be fed either gravimetrically or using a screw feeder.[25] The propensity for the production of fines can be greater in roller compaction, compared with wet granulation, depending on the properties of the powder. Ribbon density and ribbon density variability can be used to evaluate the impact of process parameters (eg, feed rate, roll speed, and roll gap) on performance. If the ribbon density is too high, the process can yield tablets with both low hardness and high friability, while if the ribbon density is too low, the produced granules can possess a wide particle size distribution that results in tablets with increased weight variability.[36] The moisture content is another quality attribute of the produced granules because it affects the flowability, cohesively, and compressibility of the granules.[37] Real-time measurements of ribbon density, composition, and moisture content has been demonstrated[38] and can be used in multi-input, multi-output feedback control strategies that can assure that ribbon density is maintained.[39] The main difference when moving to continuous manufacturing is that the dry granulation process will be integrated with upstream and downstream operations, and thus the feed rate and other properties will be impacted by the upstream units and thus, must be monitored and managed via appropriately designed control strategy.

37.4.4 Particle size reduction

The most commonly used size reduction equipment in the industry is the conical-mill (ie, the co-mill). Co-mills use an impeller to push incoming material through a conical screen. The key design parameters for a co-mill include the screen mesh size, the size of the cone, the impeller shape, and the impeller-to-screen distance, which can be adjusted using spacers. In addition to achieving the desired particle size distribution, yield can also be a concern in particle size reduction operations because the production of a large quantity of fines could result in significant losses of potentially expensive API.[18]

Co-mills are inherently continuous equipment however, the manner in which they are operated will be somewhat different. In batch production, the material to be milled is loaded on top of the co-mill and the mill is run at a set speed until all of the material has

been processed. In a continuous process, it will be necessary to match the rate at which the mill processes material to the rate at which upstream units operate.[40]

37.4.5 Compression

The compaction of powder blends is typically achieved in a tablet press, which contains several components integrated as a single processing unit. A hopper conveys material into the tablet press. A feed frame is then used to distribute the powder or granular material into the die cavity that defines the tablet size and shape. A punch compresses the material within the die and then ejects the compressed tablet. The design of tablet presses can vary parameters depending on the tooling, the number of compression stations, and feeding method (ie, force feeding or suction filling).[18] The tooling includes the die and punch size, and geometry, which can be changed according to the desired tablet weight and shape. Tooling must be engineered to withstand the stresses associated with compression to provide satisfactory reliability and maintain tablet uniformity. Key operating parameters for a tablet press include the powder feed rate and the compression force applied to the tablets, as well as the feed-frame speed (ie, the rate of tablet production). Die-fill uniformity is often a limiting factor for setting the operating speed on a rotary press and impacts tablet weight and weight variability, while compression force affects tablet properties like hardness and density.[41]

Segregation during die filling that results in content uniformity issues can occur if the powder material possesses wide particle size distributions or density variability. The time required to fill the die can be impacted by low bulk density, poor compressibility, or the cohesiveness of the powder material, resulting in tablet weight variability and insufficient tablet hardness.[18] Variability in tablet hardness can impact product performance (ie, dissolution).[42,43] Excess amount of particle fines can lead to compression problems, such as capping.[44,45] In addition, the physiochemical properties of the active ingredient could affect the propensity for temperature-dependent polymorphic transformations during the compaction process.[46]

Tablet presses are another example of equipment that currently runs continuously. However, in order for tableting to be integrated with the upstream operations in a continuous line, the dynamics of the fill level in the feed hopper may differ from those seen in traditional batch manufacturing.[18] One operating approach is to select the turret speed of the tablet press to match the desired line rate of the process. The fill level in the hopper can then be allowed to vary during operation to serve as buffer to absorb upstream flow-rate variation.

37.4.6 Coating

The film coating is typically performed for cosmetic purposes, but the film coat is also sometimes used to modify the release of API and achieve the desired tablet performance. Film coating usually involves a process where tablets are sprayed with a pigment-containing polymer solution while being rotated in a dry air stream. The drying air removes moisture, leaving behind tablets coated by a thin film of colored polymer. In continuous film coating, tablets are simultaneously loaded at one end of a long rotating perforated cylinder to pass through a multi-gun spray zone, while coated tablets continuously exit at the opposite end. The main challenges with this design are a high degree of back-mixing at low-flow rates, which result in a large variance in the amount of coating applied to each tablet. An alternative (semi-continuous) approach is to operate two short-cycle batch coaters, loading or unloading one coater while the second coater is in operation.[13] This approach can result in less intra-batch variability in film coats, but suffers from less operational flexibility with regards to adjusting to changes in the process line rate. Both approaches may be considered when designing a process train that includes a film-coating operation.[13]

37.4.7 Emerging technologies for continuous drug product production

Emerging technologies for continuous drug production manufacturing include technologies where components, primarily API and polymer excipients, are blended and homogenized. Such technologies alleviate the potential for segregation and reduce potential issues related to the flowability of materials. One challenge if the API and excipients are blended in solution is that the crystallization of the API must be taken into account in the design of the process in order to either obtain the desired API crystal form or avoid crystallization if an amorphous dispersion is desired. The properties of the resulting blend should be designed to enable the direct production of the drug product.[13] This can be more challenging for high-dose drug products where API properties have a significant impact on blend properties and thus, there is less flexibility in tuning blend properties via adjusting the formulation. Emerging continuous technologies that take advantage of these approaches include spray drying, holt melt extrusion (HME), casting, and 3D printing.[13]

Spray drying is a well-established and widely applied technology to manufacture powders and is advantageous for products where the desired solid state of the API is amorphous. The amorphous form can be utilized to improve the bioavailability of poorly

soluble APIs because drying rates are typically fast enough to prevent crystal nucleation and growth. Annealing can be used to trigger crystallization if desired, but this may not be sufficient for APIs encapsulated in a polymer matrix because the fast drying leads to a rapid viscosity increase, kinetically trapping the API in the carrier matrix.[47] Spray drying is an inherently continuous process in which a solution is sprayed through a nozzle into a vessel where a gas, such as nitrogen, is blown in order to dry the airborne droplets. The spray-drying process consists of four basic stages: atomization of the liquid, mixing of the liquid with the drying gas, evaporation of the liquid, and separation of the dried particles from the gas. Understanding the relationship between process parameters (eg, inlet temperature, drying gas properties (humidity, flow rate), feed rate, and atomization air flow rate) and formulation properties (ie, feed composition [API, carrier, solvent], solids content in the feed, solvent type, viscosity, and surface tension of the drying solution) are crucial for the reproducible production of quality material.[47,48] In order for these droplets to dry sufficiently, commercial spray-drying equipment is often required to be quite large although small-scale systems are available for process development. The main challenge for incorporating spray drying into a continuous process is that accompanying upstream and downstream steps are not inherently continuous. Specifically, upstream continuous mixing of polymer/API/solvent where the slow dissolution dynamics of polymers can impact mixing performance, and downstream secondary drying, where typically long processing times are required to meet ICH limits for residual solvents.[13] A related technique, electrospray drying, can be used to process thermolabile molecules and with sufficiently high viscosity and conductivity in which solvent evaporation from the atomized droplets takes place using an electric field at ambient temperature and pressure.[49]

HME is a continuous process that has significant potential for the manufacturing of different dosage forms and increasing the solubility of poorly soluble drugs by forming a solid dispersion in a polymer matrix.[50] The process is well-known and commonly used in the plastics industries. HME combines several unit operations in one process to produce homogeneous strands of molten material. The material is typically fed with volumetric or LIW feeders in the intake zone of the extruder. The solids melt in the plasticization zone due to high shear forces and, to a lesser extent, due to heat transfer via the extruder barrel. HME is typically limited to manufacturing drug products with stable APIs insensitive to thermal stresses because heat is applied to the formulation during processing. Next, the melt is devolatized and subsequently forced into an extrusion die.[51] The processing parameters of feeding rate, screw speed, barrel temperature profile, and screw configuration impact degree of fill, the residence time, and heat transfer. The screw configuration, for example, will impact the degree of mixing and therefore uniformity of the solid dispersion. In addition, the energy contributions from the heated barrel and from the number of mixing screw elements combined with the RTD will impact the maximum temperature of the melted phase and any thermally related attributes.[52] Tablets are relatively simple shapes that can be directly formed when an API is mixed with a flowable excipient, such as in polymer dispersions. Injection molding is one technique that can be used, since it has been used for decades to make inexpensive plastic parts.[13]

Another technology is 3D printing, in which droplets of API formulations are printed on edible substrates, such as films or excipient tablets, via an approach such as ink-jet printing.[53,54] This allows powder handling to be replaced with more readily controllable, liquid-processing operations. Liquid formulations can include solutions, polymer melts, emulsions, or suspensions. Moreover, solid depositions can be produced with API in either a crystalline or amorphous form. This approach allows for automation and precise control over material properties, fluid and deposit temperature, drug morphology, droplet size, and dissolution rate[55] and may potentially be used for personalized medication. While this approach is most advantageous for low-dose drugs, 3D printing using more conventional multilayer deposition is also feasible for high-dose drugs and has been utilized for the commercial manufacturing of an approved drug product.[56]

Strip-film manufacturing is another emerging technology that can be naturally adopted for continuous manufacturing. Solvent casting is the most widely used manufacturing method for thin films where aqueous polymer and API solutions are combined, deaerated, and transferred to the casting station where the solution is cast into films on a release liner. The cast film is dried and can then be cut into the desired shape.[57] Utilizing nanoparticle suspensions, this method can be adopted to produce finished dosage forms for poorly soluble APIs with higher drug loading.[58] Film drying via combined conduction and convection leads to commercially feasible drying times. Material properties of the polymers (eg, viscosity) used in film production have a significant impact on content uniformity and stability of the final dosage form.[59] The cutting of the film into different sizes enables flexible or metered dosing. Films can also be rolled or compressed to form multi-layer tablets, permitting the delivery of higher doses.[60] It is anticipated that other emerging technologies will be developed as the adoption of continuous pharmaceutical manufacturing spreads.

37.4.8 Process integration

In designing a continuous manufacturing process, although the analysis and optimization of individual process equipment remain important, the primary objective is to identify and evaluate design elements that pose a potential risk to product quality for the fully integrated system. In an integrated process, individual pieces of equipment (ie, unit operations) are connected in series. In such a process, a train of multiple units, one after the next, is connected via piping to sequentially perform, for example, powder-to-tablet manufacturing without isolation of intermediates. The output of a preceding unit becomes the input of the subsequent unit, with material continuously flowing from one unit to the next.

For continuous manufacturing, understanding the dynamics of this flow of material through the process is a critical aspect of process understanding and enables material traceability (the ability to preserve and access the identity and attributes of the material throughout the process). Such an understanding of process dynamics can be obtained by characterization of RTD through a tracer experiment and/or process modeling.[23,61,62] The RTD is a probability distribution that describes the amount of time a mass or fluid element remains in a process. In tracer experiments, a pulse or step change of tracer is added to the inlet of the continuous equipment, and the response of the tracer concentration profile at the outlet is measured. The overall RTD for the manufacturing process can be computed from the experiments for each unit operation measured separately, using a convolution integral.

The RTD curve can be utilized to predict the propagation of material or disturbances through the system or, in a retrospective analysis, to determine when the ingredients in a given product unit were fed to the manufacturing system. The RTD is dependent upon several factors, such as processing time, equipment parameters, and material properties and thus could be purposefully managed through active control of these process variables. One option to describe the material traveling through the system is to define a traceability resource unit (TRU).[63] A TRU can be specified as a segment of material that flows through the process together and can then serve as a unique identifier from a process history perspective to achieve traceability throughout the integrated continuous process. Understanding of process dynamics has implications on process monitoring and control, as described next.

37.4.9 Process monitoring and control

The control strategy for a continuous process should be designed to control the quality of the product in response to potential variations in the process, equipment conditions, incoming raw materials, or environmental factors over time. Due to the absence of isolated intermediates and the typically faster process dynamics, a continuous process may necessitate more frequent measurements, with real-time monitoring of process parameters and quality attributes of in-process materials, typically constituting an essential component of a control strategy for a continuous process. The adoption of an active process control system, enabled by the availability of real-time process and quality data, provides an opportunity to increase efficiency, lower operating costs, and improve control of product quality.[4] Specifically, an active process control system allows process parameters to be adjusted in response to disturbances to ensure that the quality attributes consistently conform to the established acceptance criteria. The active process control system for integrated processes should be able to appropriately respond to both fast (within a unit operation) and slow (propagating from upstream unit operations) disturbances.[64,65] The risk of producing out-of-specification product is lowered through the implementation of adaptive engineering controls.

PAT tools can be applied to measuring surrogates for the quality attributes of a final product, some of which may have already been incorporated into the control strategy for process monitoring and control. For this reason, continuous manufacturing naturally lends itself to real-time release testing (RTRT), which is the ability to evaluate and ensure the quality of in-process materials and/or final product based on process data that typically include a valid combination of measured raw material attributes and process controls.[66] A supervisory control and data acquisition system can be implemented. Such a system incorporates measurements of process parameters, incoming raw material, and in-process material attributes, as well as final product quality attributes with a model of the process to reconcile the data in order to support RTRT.[65,67] Due to a high frequency of data collection, statistical methods for large sample sizes can be applied to increase the confidence level that the batch conforms to the desired quality.[68] RTRT batch calculations should consider the observed variance in critical quality attributes over the production run to account for intra-batch variability. A risk analysis aids in consideration of PAT failure, and procedures can be developed in order to establish contingencies for process monitoring and batch release. The procedures could include end-product testing or utilizing surrogate measurements to ensure that the product maintains an acceptable level of quality.[15]

In addition to naturally lending itself to RTRT, the increase in the amount of process and quality data

collected during a continuous production facilitates the adoption of multivariate process monitoring approaches. Multivariate statistical process control (MSPC) is a process monitoring approach used to determine whether the variability in the process is stable over time.[69] It can be used to detect abnormal events in the process that may lead to adverse consequences (eg, out-of-specification product, equipment malfunction, or process safety incident) and provide diagnostic information for the event and identify points for mitigation. Taking advantage of the fact that process variables are often correlated, MSPC simplifies process monitoring by reducing the number of control charts being tracked without losing information. MSPC may also enhance the detection of abnormal process operations by identifying changes in the relationships among process parameters and quality attributes,[69] which may be difficult to detect using solely univariate process monitoring approaches.[70] Once an abnormal event is detected, then diagnosis of its cause may be facilitated by the use of the direct real-time measurement data and previously recorded abnormal event signatures to diagnose the fault and initiate appropriate mitigation actions.[71] The detection of the abnormal event can also serve as trigger for tracking the nonconforming material, which will be generated until the process returns to normal operation. Suitable strategies for segregating this nonconforming material will also be required.

37.5 CHALLENGES IN IMPLEMENTING CONTINUOUS MANUFACTURING

There are several inherent technical challenges to continuous manufacturing, including material handling, PAT, process flexibility, process models for material tracking, and optimizing process start-ups and shutdowns. The extent of these challenges will depend on the specifics of the product and manufacturing route. In continuous manufacturing, materials must continuously flow between unit operations over an extended period of time. This can be a particular challenge for powder materials, where many material attributes (eg, density, cohesion, and electrostatics) may affect flowability. The potential build-up of materials within the process over time due to adhesion presents another challenge for long manufacturing campaigns.[72]

As previously noted, real-time monitoring typically constitutes an essential component of a control strategy. The sampling interface for these PAT measurements in a continuous manufacturing system can be challenging, especially for heterogeneous powder systems. Industrial experience indicates that poor measurement performance is often attributable to sampling system issues rather than the process analyzer itself.[73] On-line and in-line measurements may reduce, but do not necessarily eliminate sampling errors.[74] Thus, sampling considerations should be assessed. For example, the location of the sensor should be evaluated to achieve representative sampling and minimize the effect of the probe on the process. Powders and dispersions limit the penetration depth of spectroscopic techniques. This may increase the importance of the sample probe location,[75] size of the sampling spot, intensity of the incident signal, etc. The sample size for the measurement should be representative of a unit dose and consider factors such as flow rate, penetration depth, and the number of scans. It is important to utilize the knowledge of the process dynamics (eg, RTD) for determining the adequate sampling frequency for PAT measurements. The measurement frequency implemented should provide sufficient resolution for the detection of a pulse of variability expected from a process disturbance.

Flexibility can be designed into a continuous manufacturing line to enable production of different drug products and/or different volumes. The first key step for incorporating flexibility is to build the process utilizing modular unit operations. Modular unit operations allow the equipment to be connected in a different order, depending on the process design (eg, the output of the blender in Fig. 37.2 can be disconnected from the input of the fluid-bed wet granulator and connected to the input of the roller compactor). Then the manufacturing line can adopt different manufacturing routes for a product through various modules, according to the formulation chosen. If a manufacturing line is design to produce multiple products, then cleanability and set-up time become extremely important.[13] Process models can facilitate the examination of different manufacturing routes. Continuous processing, compared with batch processing, offers a greater opportunity to develop and better utilize process models to gain process knowledge because their governing equations generally can be simplified. From a design standpoint, process models can be utilized as a tool to evaluate equipment configurations and manufacturing routes to determine a process configuration and ranges of unit operating parameters that will meet the desired objectives. This information can then lead to more focused experimental development efforts, which could eventually result in a higher level of process understanding. The reduction in experimentation due to process design evaluation reduces materials usage (eg, API, excipients), waste, development time, cost, and personnel exposure.[76,77]

For a given process design, process modeling can be utilized to perform sensitivity analysis to identify the

key interactions or relationships among process parameters and material or product attributes to support a quality risk assessment.[76] In addition, predictive process models can be used to understand how disturbances might propagate through a continuous process, which is another area of concern. Specifically, if a disturbance enters the first unit operation in a series, the question arises as to whether that disturbance will spread out, and if so, by how much. The RTD for each unit operation can be utilized to predict the propagation of material or disturbances through the system in order to answer this question. In addition to understanding each unit operation's transfer function, care must be taken to understand the system capacitance and delay times associated with material transport operations between steps. Care must be taken during design and testing of the individual and overall system controls to ensure that the system dynamics do not become unstable and lead to control system runaway or other chaotic phenomena because complex interconnected systems with time delays are inherently nonlinear. The careful design of mixing steps and buffer tanks in a continuous process can be used to help smooth process dynamics by dampening process variability. However, the dampening action also must be balanced against injecting sluggishness into the process dynamics, which can counteract effective control action. It is expected that process modeling can enable testing of system dynamics and control strategies offline and that model-based control of the process can be a key element in minimizing the impact of disturbances on product quality.[4,13] The development of modular equipment and the advancement of process modeling for pharmaceutical operations can enable flexible processing lines, matching the flexibility offered by batch processing.

During planned start-up and shutdown, there may be periods of time when the in-process material or product does not meet the target quality attributes. In cases where a continuous manufacturing line is producing multiple products, each manufacturing run may be a couple of days or even less than a single day, depending upon the desired production volume. Optimizing start-up and shut down then to minimize product waste is another key challenge in order to make it cost effective. Dynamic process models can be very effectively used to study and optimize start-up and shutdown strategies by taking the process through a series of well-defined intermediate states, including timely sequencing of unit operation start and stop times to decrease material losses. For example, a continuous blender could be filled completely, and mixed in a batch mode (with no discharge) for a few moments prior to allowing material to pass to the next process step. Although the process may not be at steady state, the material could still be processed so as to meet the desired quality attributes (eg, blend uniformity).[78] Utilizing available dynamic optimization methodologies can be an effective way of identifying start-up/shut-down strategies that may be challenging to discover by trial and error.[79] Reduction in wasted material can also be achieved by design by minimizing system volume to the extent possible, because the average system residence time for a given throughput will increase proportionally to the hold-up volume, and typically 3–5 residence times must pass before a process achieves steady state.

37.6 CONCLUSION

Continuous pharmaceutical manufacturing offers potential flexibility, quality, and economic advantages over batch processing, both in process development and manufacturing for the pharmaceutical sector. Over the past decade, there have been significant advancements in science and engineering to support the implementation of continuous pharmaceutical manufacturing. These investments, along with the adoption of the QbD paradigm for pharmaceutical development and the advancement of PAT for designing, analyzing, and controlling manufacturing, have progressed the scientific and regulatory readiness for continuous manufacturing. Building on this progress, research efforts should continue in several key areas to address the remaining implementation challenges.

References

1. U.S. Food and Drug Administration. *Drug recalls*. [cited March 25, 2015]; Available from: http://www.fda.gov/Drugs/drugsafety/DrugRecalls/default.htm.
2. *Strategic Plan for Preventing and Mitigating Drug Shortages*; 2013.
3. Lee SL, et al. Modernizing pharmaceutical manufacturing: from batch to continuous production. *J Pharm Innov* 2015;**10**(3):191–9.
4. Myerson AS, et al. Control systems engineering in continuous pharmaceutical manufacturing. May 20–21, 2014 continuous manufacturing symposium. *J Pharm Sci* 2015;**104**(3):832–9.
5. *Stericycle Recall Index Q4 2014*; 2015.
6. U.S. Food and Drug Administration. *Pharmaceutical CGMPs for the 21st century—A risk based approach*; 2004.
7. U.S. Food and Drug Administration. *Guidance for industry: Q7A good manufacturing practice guidance for active pharmaceutical ingredients*; 2001.
8. Schaber SD, et al. Economic analysis of integrated continuous and batch pharmaceutical manufacturing: a case study. *Ind Eng Chem Res* 2011;**50**(17):10083–92.
9. Lutze P. PSE tools for process intensification. In: Krist JKH, Gernaey V, Rafiqul G, editors. *Computer aided chemical engineering*. Elsevier; 2015. p. 35–40.
10. Poechlauer P, et al. Continuous processing in the manufacture of active pharmaceutical ingredients and finished dosage forms: an industry perspective. *Org Process Res Dev* 2012;**16**(10):1586–90.

11. Chen J, et al. Pharmaceutical crystallization. *Cryst Growth Des* 2011;**11**(4):887–95.
12. Muzzio, F. *QbD for continuous manufacturing: CQAs, CMAs, CPPs*. In: *ERC-SOPS presentation*. San Juan, PR; 2014.
13. Byrn S, et al. Achieving continuous manufacturing for final dosage formation: challenges and how to meet them. May 20–21, 2014 continuous manufacturing symposium. *J Pharm Sci* 2015;**104**(3):792–802.
14. Kuehn SE. *Pfizer's continuous manufacturing pod comes in for a landing*. 2015 [cited October 31, 2015]; Available from: http://www.pharmamanufacturing.com/articles/2015/pfizers-continuous-mfg-pod/.
15. Allison G, et al. Regulatory and quality considerations for continuous manufacturing. May 20–21, 2014 continuous manufacturing symposium. *J Pharm Sci* 2015;**104**(3):803–12.
16. Mascia S, et al. End-to-end continuous manufacturing of pharmaceuticals: integrated synthesis, purification, and final dosage formation. *Angew Chem Int Ed* 2013;**52**(47):12359–63.
17. Malet-Sanz L, Susanne F. Continuous flow synthesis. A pharma perspective. *J Med Chem* 2012;**55**(9):4062–98.
18. Rogers AJ, Hashemi A, Ierapetritou MG. Modeling of particulate processes for the continuous manufacture of solid-based pharmaceutical dosage forms. *Processes* 2013;**1**(2):67–127.
19. Engisch WE, Muzzio FJ. Feedrate deviations caused by hopper refill of loss-in-weight feeders. *Powder Technol* 2015;**283**:389–400.
20. Engisch WE, Muzzio FJ. Loss-in-weight feeding trials case study: pharmaceutical formulation. *J Pharm Innov* 2015;**10**(1):56–75.
21. Hopkins M. Loss in weight feeder systems. *Meas Control* 2006;**39**(8):237–40.
22. Vanarase AU, Osorio JG, Muzzio FJ. Effects of powder flow properties and shear environment on the performance of continuous mixing of pharmaceutical powders. *Powder Technol* 2013;**246**:63–72.
23. Gao Y, et al. Characterizing continuous powder mixing using residence time distribution. *Chem Eng Sci* 2011;**66**(3):417–25.
24. Gao YJ, Ierapetritou M, Muzzio F. Periodic section modeling of convective continuous powder mixing processes. *AICHE J* 2012;**58**:69–78.
25. Vervaet C, Remon JP. Continuous granulation in the pharmaceutical industry. *Chem Eng Sci* 2005;**60**:3949–57.
26. Dhenge RM, et al. Twin screw granulation: steps in granule growth. *Int J Pharm* 2012;**438**(1–2):20–32.
27. Vercruysse J, et al. Visualization and understanding of the granulation liquid mixing and distribution during continuous twin screw granulation using NIR chemical imaging. *Eur J Pharm Biopharm* 2014;**86**(3):383–92.
28. El Hagrasy AS, Litster JD. Granulation rate processes in the kneading elements of a twin screw granulator. *AICHE J* 2013;**59**(11):4100–15.
29. Bouffard J, Bertrand F, Chaouki J. A multiscale model for the simulation of granulation in rotor-based equipment. *Chem Eng Sci* 2012;**114**:116–21.
30. Sayin R, El Hagrasy AS, Litster JD. Distributive mixing elements: towards improved granule attributes from a twin screw granulation process. *Chem Eng Sci* 2015;**125**:165–75.
31. Vercruysse J, et al. Continuous twin screw granulation: influence of process variables on granule and tablet quality. *Eur J Pharm Biopharm* 2012;**82**(1):205–11.
32. El Hagrasy AS, et al. Twin screw wet granulation: influence of formulation parameters on granule properties and growth behavior. *Powder Technol* 2013;**238**:108–15.
33. Dhenge RM, et al. Twin screw wet granulation: effects of properties of granulation liquid. *Powder Technol* 2012;**229**:126–36.
34. Tu WD, Ingram A, Seville J. Regime map development for continuous twin screw granulation. *Chem Eng Sci* 2013;**87**:315–26.
35. Palzer S. Drying of wet agglomerates in a continuous fluid bed: influence of residence time, air temperature and air-flow rate on the drying kinetics and the amount of oversize particles. *Chem Eng Sci* 2007;**62**(1–2):463–70.
36. Akseli I, et al. A quantitative correlation of the effect of density distributions in roller—compacted ribbons on the mechanical properties of tablets using ultrasonics and X-ray tomography. *AAPS Pharmscitech* 2011;**12**(3):834–53.
37. Austin J, et al. The use of near-infrared and microwave resonance sensing to monitor a continuous roller compaction process. *J Pharm Sci* 2013;**102**(6):1895–904.
38. Gupta A, et al. A novel microwave sensor for real-time online monitoring of roll compacts of pharmaceutical powders online—a comparative case study with NIR. *J Pharm Sci* 2015;**104**(5):1787–94.
39. Hsu S-H, Reklaitis GV, Venkatasubramania V. Modeling and control of roller compaction for pharmaceutical manufacturing. *J Pharm Innov* 2010;**5**(1–2):24–36.
40. Reynolds GK. Modelling of pharmaceutical granule size reduction in a conical screen mill. *Chem Eng J* 2010;**164**(2):383–92.
41. Jackson S, Sinka I, Cocks A. The effect of suction during die fill on a rotary tablet press. *Eur J Pharm Biopharm* 2007;**65**(2):253–6.
42. Gentis ND, Betz G. Compressibility of binary powder formulations: investigation and evaluation with compaction equations. *J Pharm Sci* 2012;**101**:777–93.
43. Patel S, Kaushal AM, Bansal AK. Effect of particle size and compression force on compaction behavior and derived mathematical parameters of compressibility. *Pharm Res* 2007;**24**:111–24.
44. Wu CY, et al. Modelling the mechanical behavior of pharmaceutical powders during compaction. *Powder Technol* 2005;**152**:107–17.
45. Podczeck F. Methods for the practical determination of the mechanical strength of tablets—from empiricism to science. *Int J Pharm* 2012;**436**:214–32.
46. Zavaliangos A, et al. Temperature evolution during compaction of pharmaceutical powders. *J Pharm Sci* 2008;**97**:3291–304.
47. Paudel A, et al. Manufacturing of solid dispersions of poorly water soluble drugs by spray drying: formulation and process considerations. *Int J Pharm* 2013;**453**(1):253–84.
48. Baldinger A, et al. Quality by design approach in the optimization of the spray-drying process. *Pharm Dev Technol* 2012;**17**(4):389–97.
49. Wang M, et al. Production and characterization of carbamazepine nanocrystals by electrospraying for continuous pharmaceutical manufacturing. *J Pharm Sci* 2012;**101**(3):1178–88.
50. Roblegg E, et al. Development of sustained-release lipophilic calcium stearate pellets via hot melt extrusion. *Eur J Pharm Biopharm* 2011;**79**(3):635–45.
51. Rauwendaal C, Noriega P. *Troubleshooting the extrusion process*. Cincinnati, OH: Hanser; 2001.
52. Schenck L, et al. *Achieving a hot melt extrusion design space for the production of solid solutions. Chemical engineering in the pharmaceutical industry: R&D to manufacturing*. John Wiley & Sons Ltd; 2011. p. 819–36..
53. Hirshfield L, et al. Dropwise additive manufacturing of pharmaceutical products for solvent-based dosage forms. *J Pharm Sci* 2014;**103**(2):496–506.
54. Içten E, et al. Dropwise additive manufacturing of pharmaceutical products for melt-based dosage forms. *J Pharm Sci* 2015;**104**(5):1641–9.
55. Içten E, Nagy ZK, Reklaitis GV. Process control of a dropwise additive manufacturing system for pharmaceuticals using polynomial chaos expansion based surrogate model. *Comput Chem Eng* 2015;**83**:221–31.

56. Khaled SA, et al. Desktop 3D printing of controlled release pharmaceutical bilayer tablets. *Int J Pharm* 2014;**461**(1):105–11.
57. Gupte A, Kathpalia H. An introduction to fast dissolving oral thin film drug delivery systems: a review. *Curr Drug Deliv* 2013;**10**:667–84.
58. Sievens-Figueroa L, et al. Preparation and characterization of hydroxypropyl methyl cellulose films containing stable BCS Class II drug nanoparticles for pharmaceutical applications. *Int J Pharm* 2012;**423**(2):496–508.
59. Krull SM, et al. Polymer strip films as a robust, surfactant-free platform for delivery of BCS Class II drug nanoparticles. *Int J Pharm* 2015;**489**(1–2):45–57.
60. Mesbah A, et al. Nonlinear model-based control of thin-film drying for continuous pharmaceutical manufacturing. *Ind Eng Chem Res* 2014;**53**(18):7447–60.
61. Kvarnstrom B, Bergquist B. Improving traceability in continuous processes using flow simulations. *Prod Plann Control* 2012;**23**: 396–404.
62. Reitz E, et al. Residence time modeling of hot melt extrusion processes. *Eur J Pharm Biopharm* 2013;**85**:1200–5.
63. Moe T. Perspectives on traceability in food manufacture. *Trends in Food Sci Tech* 1998;**9**:211–14.
64. Lakerveld R, et al. Model-based design of a plant-wide control strategy for a continuous pharmaceutical plant. *AICHE J* 2013;**59** (10):3671–85.
65. Singh R, Ierapetritou M, Ramachandran R. System-wide hybrid MPC–PID control of a continuous pharmaceutical tablet manufacturing process via direct compaction. *Eur J Pharm Biopharm* 2013;**85**(3):1164–82.
66. *Guidance for Industry: Q8(R2) Pharmaceutical Development*; 2009.
67. Markl D, et al. Supervisory control system for monitoring a pharmaceutical hot melt extrusion process. *AAPS PharmSciTech* 2013;**14**(3):1034–44.
68. Edward G, Schilling DVN. *Acceptance sampling in quality control. 2nd ed.* Boca Raton, FL: CRC Press; 2009.
69. Kourti T. The process analytical technology initiative and multivariate process analysis, monitoring and control. *Anal Bioanal Chem* 2006;**384**(5):1043–8.
70. Kourti T. Application of latent variable methods to process control and multivariate statistical process control in industry. *Int J Adaptive Control Signal Process* 2005;**19**(4):213–46.
71. Gupta A, et al. Intelligent alarm management applied to continuous pharmaceutical tablet manufacturing: an integrated approach. *Ind Eng Chem Res* 2013;**52**(35):12357–68.
72. Chou, H.L., K.C. Yeh, *Method of controlling adhesive build-up on a yankee dryer*.Google Patents. 2010.
73. Waters T. *Industrial sampling systems: reliable design & maintenance for process analyzers*. Solon, OH: Swagelok Company; 2013.
74. Vanarase AU, et al. Development of a methodology to estimate error in the on-line measurements of blend uniformity in a continuous powder mixing process. *Powder Technol* 2013;**241**:263–71.
75. Wahl PR, et al. Inline monitoring and a PAT strategy for pharmaceutical hot melt extrusion. *Int J Pharm* 2013;**455**(1):159–68.
76. Boukouvala F, et al. An integrated approach for dynamic flowsheet modeling and sensitivity analysis of a continuous tablet manufacturing process. *Comput Chem Eng* 2012;**42**(0): 30–47.
77. Sebastian Escotet-Espinoza M, Maitraye Sen RS, O'Connor T, Lee S, Chatterjee S, Ramachandran R, et al. Flowsheet models modernize pharmaceutical manufacturing design and risk assessment. *Pharm Technol* 2015;**April**:34–42.
78. Benyahia B, Lakerveld R, Barton PI. A plant-wide dynamic model of a continuous pharmaceutical process. *Ind Eng Chem Res* 2012;**51**:15393–412.
79. Giridhar A, et al. *Intelligent process management for continuous operations in pharmaceutical manufacturing*. 24th European symposium on computer aided process engineering. Elsevier; 2014.

PART IV

REGULATORY ASPECTS OF PRODUCT DEVELOPMENT

This page intentionally left blank

CHAPTER

38

Drug Product Approval in the United States and International Harmonization

L. Wu[1], *H. Smith*[1], *H. Zheng*[2] *and L.X. Yu*[1]

[1]U.S. Food and Drug Administration, Silver Spring, MD, United States [2]Albany College of Pharmacy and Health Sciences, Albany, NY, United States

38.1 DRUG PRODUCT APPROVAL AND THE US FOOD AND DRUG ADMINISTRATION

Any person or organization who intends to bring a drug product to market in the United States must, by law, first submit a marketing application to the US Food and Drug Administration (FDA). The FDA is responsible for reviewing each marketing application and determining whether the applicant or sponsor has provided adequate evidence of safety, efficacy, and quality to warrant marketing approval. FDA oversight of drug products extends, as described in Chapter 21 of the Code of Federal Regulations and guidance documents, over the lifecycle of marketed drug products and to all relevant drug manufacturing activities and facilities. This chapter presents a current overview of the drug approval process as it pertains to new drug applications (NDAs), abbreviated new drug applications (ANDAs), and biologics license applications (BLAs) in the context of FDA.

38.1.1 History and background of drug regulations in the United States

The US FDA is a scientific, regulatory, and public health organization, and is the oldest federal agency dedicated to consumer protection. The history of the US federal food and drug regulatory system can be traced back to the commercial and custom food laws of the 1800s. The Biologics Control Act of 1902 introduced premarket requirements for some biologic therapeutic products. The Pure Food and Drug Act of 1906 was the first comprehensive legislation for foods and drugs, identifying official standards for drugs and offering definitions for misbranding and adulteration; it was enforced by the US Department of Agriculture (USDA) Bureau of Chemistry. In 1927, the Bureau of Chemistry was reorganized within the USDA and gave rise to the Food, Drug, and Insecticide Administration, which was renamed the FDA 3 years later and moved from the USDA to become the forerunner of today's Department of Health and Human Services. The Food, Drug, and Cosmetic Act of 1938 (FDCA) was a complete overhaul of the 1906 Act and gave the FDA responsibility for approving the marketing of new drug products on the basis of safety considerations. Criteria for drug efficacy were added to the approval process with the 1962 Kefauver-Harris amendments to the FDCA.

In the 1980s, The Drug Price Competition and Patent Term Restoration Act of 1984, commonly known as the Hatch-Waxman Act, was designed to promote competitiveness in the development of innovative drug products and to bring more affordable versions of approved drugs—that is, generic versions—to the market.[1] The Hatch-Waxman Act created the ANDA process, whereby applicants of generic drug products may generally rely on the clinical studies used to support approval of an innovator product (ie, the reference listed product (RLD)) from a safety and efficacy perspective. The generic drug approval process thus avoids much of the time and cost associated with the development of "innovator" drugs. Moreover, the Public Health Service (PHS) Act of 1944 introduced oversight of biologics, whereas its amendment under section 351(k), signed into law on Mar. 23, 2010, created an abbreviated licensure pathway under the Biologics Price Competition and Innovation Act (BPCI Act) for biological products that are demonstrated to be

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00038-8

"biosimilar" to or "interchangeable" with a biological product already licensed under section 351(a) of the PHS Act.[2]

More recent legislation that pertains to drug regulation by the FDA includes the FDA Safety and Innovation Act of 2012 (FDASIA).[3] Under FDASIA, the FDA has been authorized for the first time to collect user fees upon receipt of applications for the marketing of generic and biosimilar drug products; FDASIA also reauthorized, for the fifth time since 1992, the collection of prescription drug user fees. The user fee acts (UFAs) for prescription drugs (PDUFA), generic drugs (GDUFA), and biosimilars (BsUFA) commit the agency to the timely review of marketing applications and allow the agency to increase the staff and resources dedicated to this purpose.

38.1.2 Current organization of the FDA

The FDA consists of the Office of the Commissioner and four directorates—Office of Foods and Veterinary Medicine, Office of Global Regulatory Operations and Policy, Office of Medical Products and Tobacco, and Office of Operations—each including substantive centers and offices, responsible for overseeing the core functions of the agency. The Office of the Commissioner provides leadership for the FDA's science and policy initiatives and ensures consistent program direction and management across the agency. The FDA Office of the Chief Counsel provides legal services to FDA on matters involving FDA-regulated drugs, reviews and evaluates FDA enforcement actions, and helps defend the agency in litigations.

Within the four directorates, six substantive centers (Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), Center for Veterinary Medicine, Center for Food Safety and Applied Nutrition, Center for Tobacco Products, and Center for Devices and Radiological Health (CDRH)) and the Office of Regulatory Affairs (ORA) conduct the bulk of FDA's scientific and regulatory evaluations and apply statutory authority over drugs (human and veterinary), food, cosmetics, tobacco, and devices and radiological products.[4]

ORA is the lead office for all agency field activities, bearing the main responsibility for inspecting establishments that produce FDA-regulated products, including foods and veterinary medicine, medical products, and tobacco and tobacco-related products. ORA professionals work collaboratively with the six FDA centers to protect patients and consumers and provide public outreach and education.[5] ORA and CDER collaborate together to ensure quality of drugs for human use as relates to their manufacture.

38.1.3 Center for Drug Evaluation and Research (CDER) organization

CDER is responsible for the timely review of all data submitted by sponsors relevant to the safety, quality, and efficacy of drug products (which include prescription, generic, over-the-counter (OTC), and selected biologic drug products). The CDER organizational chart is shown in Fig. 38.1. A set of review

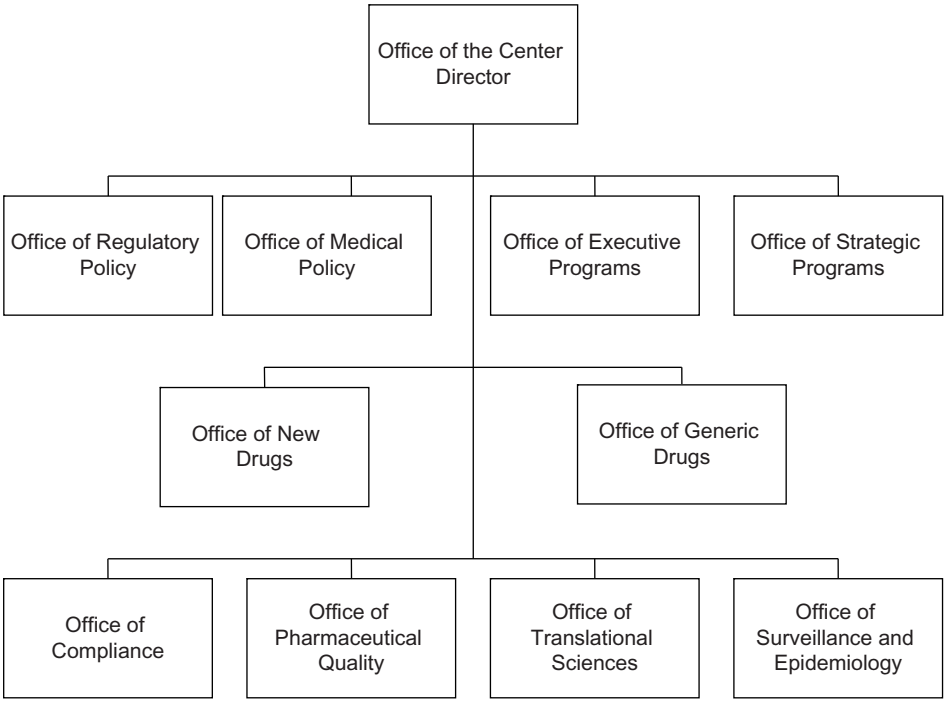


FIGURE 38.1 The organizational structure of the FDA Center for Drug Evaluation and Research.

initiatives taken at CDER, including 21st Century Review, Computational Science Center, Critical Path, Equal Voice, and Pharmaceutical Quality for the 21st Century, are essential for a current understanding of drug product evaluations within CDER.[6] In addition, the launch of Office of Pharmaceutical Quality (OPQ) in 2015 has greatly reconfigured collaborative structures within CDER as well as between CDER and ORA, strengthening the agency's focus on pharmaceutical quality, as we discuss next.

38.1.4 Pharmaceutical quality oversight[7]

The chemistry, manufacturing, and controls (CMC) section of marketing applications, wherein information is presented that is essential to the FDA assessment of the identity, strength, quality, and purity of drug substances and products, is under the purview of OPQ. OPQ is strategically organized to streamline regulatory processes, advance regulatory standards, align areas of expertise, and originate surveillance of drug quality. OPQ intends to set the global benchmark for regulation of pharmaceutical quality across all drug products (new, generic, OTC) and across all sites of manufacturing (domestic and foreign). The launch of OPQ within CDER integrates the review, inspection, surveillance, policy, and research of all drug substances (also known as active pharmaceutical ingredients (APIs)) and drug products. OPQ orchestrates many aspects of the regulatory review of NDAs, ANDAs, and BLAs, and centralizes functions for policy-making, research and science activities, project management, quality management systems, and administrative activities. The organization of OPQ is depicted in Fig. 38.2. OPQ creates a single umbrella organization that consolidates all CDER quality review by establishing its Office of New Drug Products (ONDP), Office of Lifecycle Drug Products (OLDP; for generic drug products and postapproval activities of new molecular entities (NMEs) and non-NMEs, 3 years and 1 year after approval, respectively), and Office of Biotechnology Products (OBP; for biologics and biosimilars). Process, facility, and microbiology reviewers within the OPQ Office of Process and Facilities (OPF) evaluate pharmaceutical manufacturing process design and controls and assess the feasibility of implementation at a commercial scale, participating in facility inspections as determined through risk- and science-based principles. The Office of Program and Regulatory Operations (OPRO) is accountable for leading and coordinating regulatory operations, maintaining a quality management system, and driving professional development and learning programs in collaboration with OPQ review offices. The Office of Policy for Pharmaceutical Quality (OPPQ) develops, implements, and updates science- and risk-based policies, standards, and guidance documents. The OPQ Office of Surveillance (OS) maintains information on all facilities involved in the manufacture of drugs destined for US patients and conducts risk analysis and monitoring across the entire inventory of facilities. The OPQ drug product quality laboratories (including the Office of Testing and Research (OTR) and the OBP laboratory component) conduct research to support the development of scientific standards and policies and provide

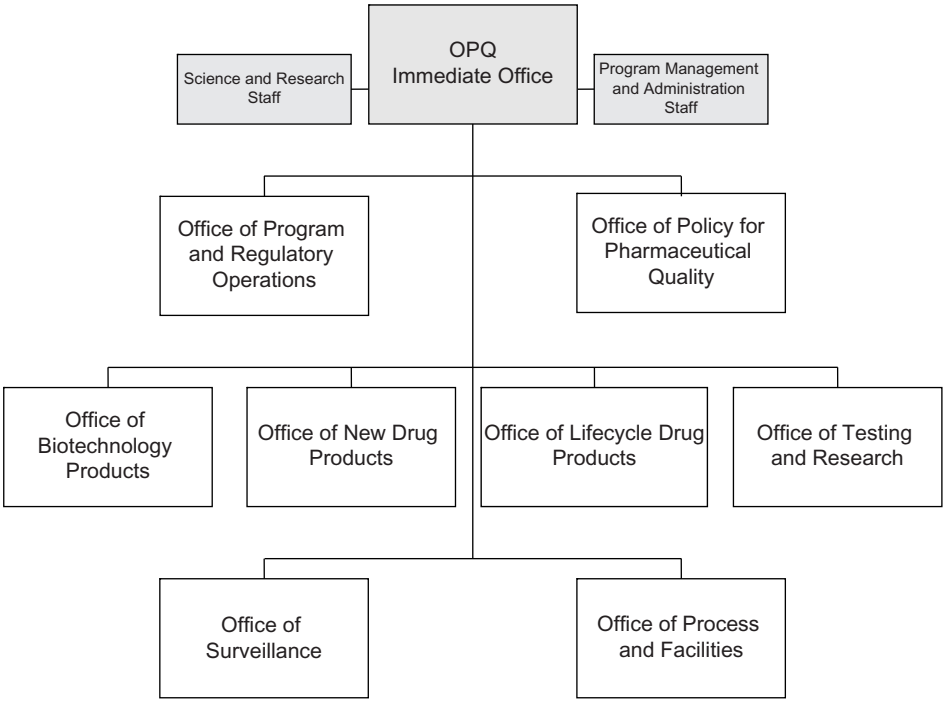


FIGURE 38.2 The organizational structure of the CDER Office of Pharmaceutical Quality.

advice, collaborative research opportunities, and scientific training for review staff. The OPQ Immediate Office consists of administrative staff and a science and research staff committed to emerging technologies and the elaboration of knowledge management within OPQ.

38.1.5 New review initiatives

As outlined in the goals for Pharmaceutical Quality for the 21st Century,[8] OPQ makes use of new review initiatives to encourage implementation of risk-based approaches that focus industry and FDA attention on the critical areas to facilitate modern quality management techniques, encourage early adoption of new technology advances, and ensure that regulatory review, compliance, and inspection policies are based on state-of-the-art pharmaceutical sciences. With the ultimate goal of putting patients first by balancing risk with availability, current initiatives in OPQ integrate review and inspection functions, under the motto of "One Quality Voice," by utilizing and developing staff expertise, facilitating continuous improvement of current review practices, and implementing question-based review (QbR) as a tool for risk-based review.

38.1.5.1 Team-based integrated quality assessment

The foundation of OPQ was conceptualized on a new review initiative known as team-based integrated quality assessment (IQA).[9] IQA maximizes reviewer expertise and effectively aligns patient-focused, science- and risk-based drug product quality recommendations, inclusive of drug substance, drug product, manufacturing, microbiology, biopharmaceutics, and facilities (Fig. 38.3). The team-building impetus within IQA integrates reviewers and field investigators and engenders effective and efficient communication, risk analysis, and quality assessments pertaining to facilities, products, and marketing applications. Reviewers and field investigators work collaboratively and are supported by project management and technical leadership; in many cases, the application reviewers participate in inspections.

Compared with the old review model that used a single reviewer for all quality information in an application, often resulting in inconsistent communication between review disciplines, the OPQ review model employs a team of experts to achieve "One Quality Voice" through collaborative assessment and consistent communication. The structure of an IQA team is shown in Table 38.1. Review disciplines may include drug substance, drug product, process, facility, microbiology, biopharmaceutics, and ORA investigators. Team advisors may come from OPQ laboratories, policy, surveillance, and other offices as needed. The application technical lead (ATL) is responsible for overseeing the scientific content of the assessment, while the regulatory business project manager (RBPM) is responsible for process and timeline.

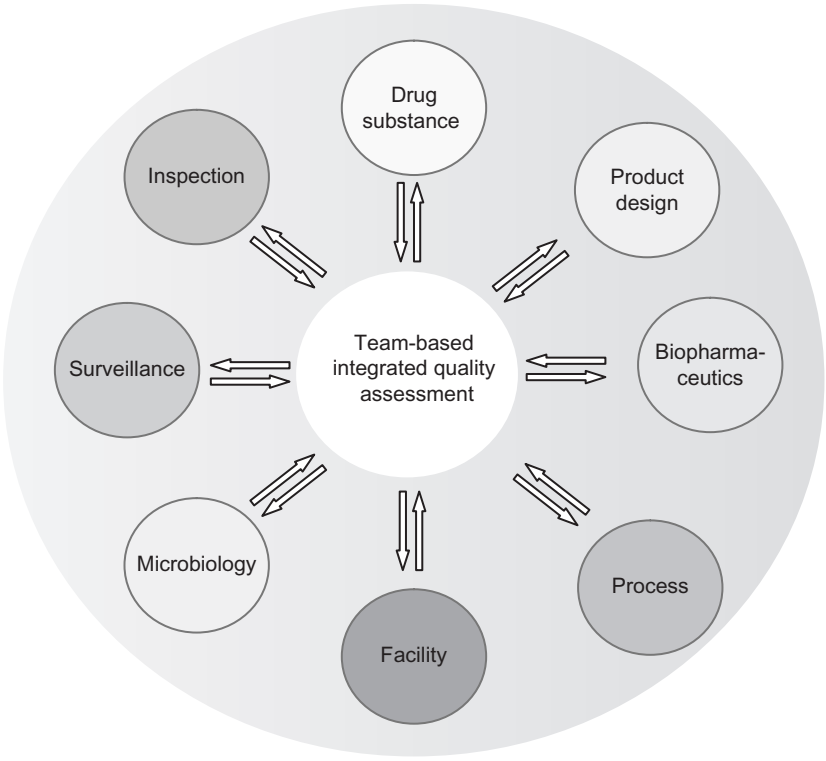


FIGURE 38.3 High-level illustration of team-based Integrated Quality Assessment.[9]

TABLE 38.1 Structure of Integrated Quality Assessment Team[10]

Role/task	Responsibility[a]
Scientific content/initial risk assessment	Application technical lead (ATL)
Process and timeline	Regulatory business project manager (RBPM)
IQA executive summary	ATL/IQA team
Assessment of drug substance	Drug substance (DS) reviewer
Assessment of drug product	Drug product (DP) reviewer
Assessment of manufacturing process	Process reviewer
Assessment of facilities	Facility reviewer
Assessment of biopharmaceutics	Biopharm reviewer
Assessment of microbiology	Microbiology reviewer
Assessment of environmental analysis	Environmental analysis (EA) reviewer
Labeling and package insert	DP reviewer
Facility inspections	ORA leads/subject matter experts (SMEs) participate
Lifecycle knowledge management	ATL/IQA team

[a]*Represents general cases.*

In a typical IQA review process, the following steps are taken: (1) the ATL performs the initial risk assessment; (2) review team is assembled; (3) timelines, preliminary findings, filing issues, etc. are discussed at the kick-off meeting; (4) there is constant communication between the review team members through formal and informal meetings and discussions; (5) consolidated quality comments are discussed, finalized, and communicated to the applicant; (6) review is finalized and the final "One Quality Voice" recommendation is communicated to CDER internal stakeholders, such as Office of New Drugs (OND) or Office of Generic Drugs (OGD). This process facilitates knowledge sharing between the team members; as members of the team, ORA investigators become apprised of any quality issue uncovered by OPQ reviewers; conversely, reviewers are updated as to any inspection findings. The collaboration between the OPQ reviewers and ORA field investigators ultimately enhances the quality assessment of the product and leads to more appropriate decisions on approvability of applications and acceptance of facilities. Moreover, by employing the IQA approach, timeframes are met effectively when aligned with user fee programs deliverables (mid-cycle, late cycle, etc.).

38.1.5.2 *Question-based review (QbR)*

QbR was fully implemented in 2007 as a tool for evaluation of CMC/quality information in ANDAs and has since been implemented as a review tool for the other applications. QbR questions were revised in 2012 and 2014 to include quality-by-design (QbD) concepts; new QbR templates for both NDA and ANDA are currently being developed that would support the integrated team-based review approach within OPQ,[11] and the trend toward using QbR will likely widen. The QbR approach transforms the CMC review into a modern, science- and risk-based pharmaceutical quality assessment; it reflects the FDA's "21st century cGMP initiatives," incorporates established components of CDER review (eg, the CDER 2004 Clinical Pharmacology and Biopharmaceutics Review Template),[12] and resonates with other regulatory authorities (eg, Health Canada). QbR is intended to guide the industry by making the FDA's CMC expectations clear and to facilitate reviewer evaluations of product quality. The wide-ranging benefits of QbR include: (1) the promotion of QbD principles by the sponsor across product development, leading to performance-based specifications; (2) facilitation of risk-based assessments by both the sponsor and the agency; (3) process improvements that minimize the need for supplement submissions; (4) consistency and transparency of CMC evaluations across the entire pharmaceutical industry; and (5) optimization of CMC review efficiency.[13]

The questions incorporated into QbR relate to drug substance and drug product quality standards, process understanding, and scale-up; help the reviewer to comprehensively assess critical formulation and manufacturing process variables; set regulatory specifications relevant to quality; and determine the level of risk associated with the manufacture and design of the product.

QbR questions and answers are incorporated in the quality overall summary (QOS) in Module 2 of the common technical document (CTD). The CMC data provided in Module 3 of the CTD are always reviewed in conjunction with the QbR questions and answers[14] (see discussions of QbR in chapters: "Modern Pharmaceutical Regulations: Quality Assessment for Drug Substances" and "Modern Pharmaceutical Regulations: Quality Assessment for Drug Products" of this edition).

38.1.5.3 *Emerging technologies*

Recognizing that drug shortages and product recalls are commonly related to compromised product or facility quality, OPQ's efforts include sustaining improvements in manufacturing quality. OPQ's Emerging Technology Team (ETT) encourages innovation through continuous dialogue with academia, industry, and other governmental agencies, and clarifies and establishes

regulatory policies for the introduction of new technologies, such as continuous manufacturing and 3D printing. The aim of ETT is to support advances in pharmaceutical manufacturing technology, which potentially will address drug shortages and product recalls and revigorate the pharmaceutical manufacturing sector in the United States. The knowledge gained from collaboration among OPQ and external stakeholders will eventually translate into new policies that reflect state-of-the-art manufacturing science and encourage manufacturers to adopt new manufacturing technologies in support of a state of continuous improvement. One example is the FDA's encouraging continuous pharmaceutical manufacturing; this technique offers potential flexibility, quality, and economic advantages over traditional batch-processing methods. Another example is 3D printing. In fact, the FDA has recently approved the first 3D-printed tablet in a review process that proved the efficiency of OPQ initiatives in team-based IQA, emerging technology, structured review format (QbR), integration of scientific research and review, as well as of facilities and inspection activities[15] (see discussions of emerging manufacturing technology and continuous manufacturing in chapter: "Emerging Technology for Modernizing Pharmaceutical Production: Continuous Manufacturing" of this edition).

38.1.6 Current drug approval overview

When a NDA/BLA is submitted to the FDA, it is the responsibility of the applicant or sponsor to provide evidence that the product(s) to be reviewed are safe, effective, and of high quality. The FDA has brought new elements of flexibility into the application review process over the past few decades, but the overall drug development and approval process generally follows a few defined steps: (1) preclinical investigation, (2) clinical investigation, (3) postapproval marketing surveillance, and (4) lifecycle management (Fig. 38.4). An overview of the stages of drug development, relevant application submissions, and corresponding regulatory paths is provided next.[16]

Before any clinical investigations may begin concerning an experimental drug, the sponsor must submit an investigational new drug (IND) application. Clinical investigation generally consists of three phases. In Phase I, safety studies are conducted in a small number of individuals (20−200 persons); in Phase II, efficacy studies begin in volunteers (up to several hundred) of the target population; and in Phase III, the most extensive and expensive part of drug development, human testing continues in a substantial number of patients (several hundred to several thousand).[17] In general, once Phase III is complete, with all clinical and nonclinical data collected, the sponsor can submit a NDA under section 505(b)(1) of the FDCA. In particular, applications for drugs that show great promise in phase II can also be submitted under FDCA 505(b)(1) and approved in an accelerated approval process with the caveat that the sponsor must still perform confirmatory phase III studies. In the case of a sponsor who intends to rely, to any extent, upon safety or effectiveness investigations that were not conducted directly by or for the sponsor, the NDA may be submitted in accordance with FDCA 505(b)(2). Whereas the 505(b)(1) application relies on full reports of investigations of safety and effectiveness conducted by or for the sponsor, or on investigations for which the sponsor has obtained a right of reference or use, for the 505(b)(2) application, at least some of the information required for approval comes from data not developed by the sponsor, such as published literature or the FDA's findings of safety and effectiveness of a previously approved drug product.[18]

Once approved, the drug product can be marketed; moderate or major modifications to approved applications should be submitted to the FDA for review as postapproval changes, referred to as supplements. After NDA approval, the next stage in the lifecycle of a drug product occurs when all forms of exclusivity (eg, patent protection) have expired for the "innovator," or RLD. At that point, under FDCA section 505(j), the RLD may be cited in the submission of an ANDA, through which the applicant seeks to market a generic version of the drug product. An ANDA contains information to show that the proposed generic product is pharmaceutically equivalent to the RLD. The proposed product and the RLD contain the same active ingredient, are of the same dosage form and route of administration, are identical in strength or concentration, and meet compendial or other established standards.[19] An important regulatory standard, termed "bioequivalence," provides the basis for establishing product substitutability (discussed next). Through reliance on bioequivalence (BE), ANDAs do not entail the

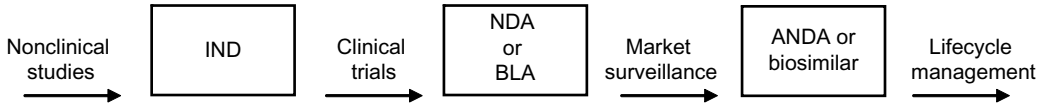


FIGURE 38.4 Typical drug development process, emphasizing the relative positions of the various drug marketing applications in the lifecycle review of drug products.

TABLE 38.2 Major Drug Approval Regulatory Pathways and Implementation Acts

Statute and CFR section	Categories	Original legislation	User fee acts
FDCA 505(b) (21CFR 314)	NDA	FDCA 1938	PDUFA
FDCA 505(b)	NDA	Waxman-Hatch 1984	PDUFA
FDCA 505(j)	ANDA	Waxman-Hatch 1984	GDUFA
PHS 351 (42USC 262)	Biologics	BCA/PHSA 1902/1944	PDUFA
PHS 351(k)	Biosimilars	BPCIA 2009	BsUFA

expenditures of time and money that normally characterize the submission of NDAs.

Whereas the NDA is used for marketing approval under FDCA provisions, the BLA can be used, under section 351 of the PHS Act, by applicants seeking to market biological products. Submission of a BLA obviates the need for submitting an NDA[20]; the BLA must contain information on product, manufacturing process, and manufacturing facilities, and licensure commits the applicant to the continued safety, purity, and potency of the product. The abbreviated pathway for a BLA is a biosimilar, established through section 351(k) of the PHS Act, according to which the proposed biosimilar must be highly similar—showing no clinically meaningful differences, in terms of safety and effectiveness—to an FDA-approved biological reference product. Only minor differences in clinically inactive components are allowable in biosimilar products.

A list of major approval regulatory pathways, along with their basis in statute is provided in Table 38.2. More details on the format and requirements for each type of submission can be found in corresponding individual sections of this chapter.

38.2 THE NEW DRUG APPLICATION PROCESS

Submitted NDAs should comprehensively provide all clinical and nonclinical drug development efforts undertaken by or on behalf of the sponsor. CDER reviewers will assess all data to conclude whether adequate evidence has been established with regard to safety, efficacy, risk-benefit profile, proposed labeling, and quality.[21]

38.2.1 INDs and presubmission of NDAs

Before conducting any of the clinical trials that will provide data for review in the NDA, sponsors are required to have submitted an IND application; the IND should summarize key attributes including appropriate stability of the drug substance, drug product, and their preparation, evidence of safety, and efficacy from preclinical studies, and should demonstrate the preparedness of investigators for clinical trials.[22] An IND may fall within two categories: commercial or research (noncommercial). A Commercial IND is an IND for which the sponsor is usually either a corporate entity or one of the institutes of the National Institutes of Health (NIH).[23] A Research IND is usually submitted by a physician for an unapproved drug or to investigate an approved drug for a new indication or new patient population (ie, "Investigator" IND) and anticipates the submission of an NDA. Additional types of INDs, such as Emergency-Use INDs and Treatment INDs, are essential in terms of public health. For patients who do not meet enrollment criteria for an existing study protocol, an Emergency-Use IND may make an investigational therapy available to a patient in a very short time frame. Treatment INDs may allow expanded access, in certain contexts of serious and life-threatening conditions, to experimental drugs that have not completed clinical testing.

Good communication between the sponsor and FDA can greatly facilitate the process that the FDA must follow in receiving and evaluating an IND submission. Sponsors can request a pre-IND meeting to receive the agency's feedback on adequate information about the investigational drug substance, drug product and their manufacture, data from animal pharmacology and toxicology studies, and the proposed clinical investigation protocols, including information about investigator to information pertaining to chemistry, manufacturing, and controls to be provided in the original IND submission. All the information must be provided in the original IND submission before the FDA allows the clinical study to proceed. The following list describes IND content; also see Table 38.3.

1. *Cover sheet*—contains sponsor and investigator information, identification of the clinical investigation phase(s), and any commitments
2. *Table of contents*
3. *Introductory statement and general investigational plan*—provides information about the dosage form and route of administration, a brief history of the drug, and a description of the overall plan for clinical investigations

TABLE 38.3 List of IND Content and Regulatory Source Documents[24]

Content (documentation)
1. Cover sheet (eg, 21 CFR 313.23(a)(2))
2. Introductory statement and general investigational plan (eg, 21 CFR 313.23(a)(3))
3. Investigator's brochure (eg, 21 CFR 313.23(a)(5))
4. Protocols (eg, 21 CFR 313.23(a)(6))
5. Chemistry, manufacturing, and control information (eg, 21 CFR 313.23(a)(7)(iv))
6. Pharmacology and toxicology information (eg, 21 CFR 313.23(a)(8))
7. Previous human experience with investigational drug (eg, 21 CFR 313.23(a)(9))
8. Additional information (eg, 21 CFR 313.23(a)(10))
9. Relevant information as requested by the FDA (eg, 21 CFR 313.23(a)(11))

4. *Investigator's brochure*—contains a brief description of the drug substance and formulation; information on the pharmaceutical and toxicological effects of drugs in animals and, if known, in humans; safety and effectiveness in humans from prior clinical studies; and anticipated possible risks and side effects
5. *Protocols*—contains detailed descriptions for each planned study, including objectives of the study, criteria for patient selection, design of study, methods used, and results
6. *Chemistry, manufacturing, and control information*—includes basic information on drug substance (eg, structure), drug product (eg, formulation), any placebo or comparator used in a controlled clinical trial, supportive stability data, labeling to be provided to each investigator, and environmental analysis requirements
7. *Pharmacology and toxicology information*—provides information from pharmacological and toxicological studies used to conclude that the drug is reasonably safe for clinical investigations
8. *Previous human experience with investigational drug*—contains description of prior experience with the investigation or marketing of the investigational drug either in the United States or in other countries
9. *Additional information*—contains information regarding drug dependence and abuse potential and description of studies performed for special drugs, such as radioactive dosimetry calculations or pediatric safety and effectiveness
10. *Relevant information*—contains any other relevant information, as requested by the FDA

The IND should demonstrate that the sponsor can adequately produce and supply consistent batches of the drug product, that the proposed product is reasonably safe for initial testing in humans, and that the clinical investigators are qualified to perform their assigned trial duties. The IND must be submitted to the FDA at least 30 days before the first-in-human (FIH) trial may be initiated. If the FDA does not respond to the IND within 30 days, the sponsor may proceed with clinical trials per the application. If the FDA has concerns about the IND or about the conduct or results of ongoing clinical trials, the agency may impose a "clinical hold." Successfully elaborated clinical development, per IND regulations, will typically emanate from completed clinical investigation Phase I, II, and III, as described in Section 38.1.6.[17] The FDA review of Phase I focuses on product safety, whereas in addition to ongoing safety evaluation, review of Phase II and III submissions includes assessment of efficacy, clinical investigations, and the likelihood of meeting the statutory standards for marketing approval. With regard to quality, FDA reviewers assess whether chemistry and manufacturing data, pertaining to drug substance and drug product, may indicate health risks to subjects enrolled in IND trials. Such risks may be related to impurities in the drug substance emanating from synthesis or subsequent degradation, physical/chemical instability of the drug product, or poorly characterized components, toxic chemical structures, or poorly characterized cell banks (master or working).[25] Information on the drug substance and drug product should be provided in a summary report, containing physical, chemical, and biological characteristics relevant to drug substance, composition of the drug product; a brief description of the manufacturing process; acceptable limits and analytical methods used to assess drug identity, strength, quality, purity, and stability. It is expected that the amount of the CMC information will vary with the phase of investigation, scope of the proposed clinical investigation, and production scale; upon submission of an NDA, however, CMC data and the quality of drug substance and drug product should be rigorously documented.

38.2.2 Format and content of the NDA

The purpose of the NDA is to supply documentation of "adequate and well-controlled investigations," including data gathered during preclinical (eg, animal, cell culture) and human clinical investigation for FDA review. The format and content of NDA submissions to the FDA are specified in regulations (21CFR 314.50) and reflected in "harmonized" application form FDA 356(h), used for all drug products (and therefore also in BLA requirements, as discussed next). The contents of an application specified in this form are listed next; also see Table 38.4.

TABLE 38.4 NDA Content and Regulatory Source Documents for Drugs and Biologics[26]

Content/item (documentation)
1. Index (Table of Contents)
2. Labeling (*check one*) _ Draft Labeling _ Final Printed Labeling
3. Summary (21 CFR 314.50 (c))
4. Chemistry Section
a. Chemistry, manufacturing, and controls information (eg, 21 CFR 314.50(d)(1); 21 CFR 601.2)
b. Samples (21 CFR 314.50 (e)(1); 21 CFR 601.2 (a)) (Submit only upon FDA's request)
c. Methods validation package (eg, 21 CFR 314.50(e)(2)(i); 21CFR 601.2)
5. Nonclinical pharmacology and toxicology section (eg, 21 CFR314.50(d)(2); 21 CFR 601.2)
6. Human pharmacokinetics and bioavailability section (eg, 21CFR 314.50(d)(3); 21 CFR 601.2)
7. Clinical microbiology section (eg, 21 CFR 314.50(d)(4))
8. Clinical data section (eg, 21 CFR 314.50(d)(5); 21 CFR 601.2)
9. Safety update report (eg, 21 CFR 314.50(d)(5)(vi)(b); 21 CFR 601.2)
10. Statistical section (eg, 21 CFR 314.50(d)(6); 21 CFR 601.2)
11. Case report tabulations (eg, 21 CFR 314.50(f)(1); 21 CFR 601.2)
12. Case report forms (eg, 21 CFR 314.50 (f)(2); 21 CFR 601.2)
13. Patent information on any patent which claims the drug (21U.S. C. 355(b) or (c))
14. Patent certification with respect to any patent which claims the drug (21 U.S.C. 355 (b)(2) or (j)(2)(A))
15. Establishment description (21 CFR Part 600, if applicable)
16. Debarment certification (FDCA 306 (k)(1))
17. Field copy certification (21 CFR 314.50 (l)(3))
18. User fee cover sheet (Form FDA 3397)
19. Financial information (21 CFR Part 54)
20. Other (Specify)

1. *Index*—A detailed index is necessary to guide FDA reviewers through the entire NDA. The index must clearly describe the contents and location of each section by volume and page number.
2. *Labeling*—The inclusion of drug product labeling is essential to the NDA and must be submitted electronically using structured product labeling (SPL) formatting in compliance with the Physician's Labeling Rule (PLR), as described at 21 CFR 201.56 and 201.57.[27]
3. *Summary*—The NDA summary provides an overview of safety, efficacy, and quality of the drug product for its proposed use. It states conclusions derived from the most important data within the NDA. Content of the summary should be presented as follows[28,29]:
   a. Proposed text of labeling for the drug—annotated;
   b. Pharmacological class, scientific rationale, intended use, and potential clinical benefits;
   c. Foreign marketing history;
   d. Chemistry, manufacturing, and controls summary;
   e. Nonclinical pharmacology and toxicology summary;
   f. Human pharmacokinetic and bioavailability summary;
   g. Microbiology summary (if required);
   h. Clinical data summary and result of statistical analysis;
   i. Discussion of benefit/risk relationship and proposed postmarketing studies.

   The summary is important because it provides FDA reviewers with an indication of the overall quality and accuracy of the NDA and the information it presents. The summary should be written as if prepared for publication in a recognized scientific journal and should integrate all salient information in the application to provide a general understanding of the drug product.
4. *CMC/pharmaceutical quality*—The pharmaceutical quality section (also known as the CMC, or chemistry section) provides detailed information on the composition, design, manufacture, test methods, specifications, container closure system, and stability of the drug substance and the final drug product.

   The drug substance subsection should detail all pertinent physical and chemical properties and tests performed to demonstrate the identity, purity, potency, and stability of the drug substance. The consistency and methods of its manufacture, along with a description of the manufacturer (eg, operations and operations control, facility layout, relevant product and processing portfolios, and prevention of contamination), should be described in adequate detail. Although drug substance information is often included in the NDA submission, information to be submitted in an application for drug substance is often submitted in a drug master file (addressed in Section 38.3.4.2 of this chapter).

   Similarly, the drug product subsection must present a rigorous account of product characterization, product and process design, manufacture and packaging, drug product specifications, microbiology, container closure system, and stability. Recommendations are provided next regarding the CMC information that should be included[30]:
   a. Description and composition of the drug product—states the name, function, and amount of each active and inactive ingredient in the drug product and describes the dosage form and container closure systems proposed for marketing

b. Pharmaceutical development report—contains information on the development studies conducted to establish that the dosage form, formulation, manufacturing process, container closure system, microbiological attributes, and usage instructions are appropriate for the purpose specified in the application; includes description of any QbD efforts to enhance product and process understanding and process control, including defining the quality target product profile (QTPP) as it relates to quality, safety, and efficacy, determining critical quality attributes (CQAs) of the drug product, linking raw material attributes and critical process parameters (CPPs) to CQAs and performing risk assessment, developing a design space, designing and implementing a control strategy, and managing product lifecycle and continual improvement[31,32]

c. Manufacture—states the names, addresses, and responsibilities of the manufacturers, describes the manufacturing and packaging processes, including flowcharts indicating production steps with equipment, materials and in-process controls and tests (eg, operating parameters, environmental controls, process tests, and in-process material tests) for the finished dosage form

d. Control of excipients—describes the excipients, lists all the specifications and tests performed to confirm their quality, and identifies novel excipients and excipients of human or animal origin

e. Control of drug product—describes specification and tests selected to assure product quality (related to safety and efficacy) and lot-to-lot consistency of the finished product; presents sampling procedures and method validation results; lists all expected drug product impurities; and identifies degradation products, residual solvents, and other miscellaneous drug product impurities

f. Reference standards—provides information on the reference standard(s) used to test the drug product or its impurities

g. Container closure systems—describes the proposed container closure system(s) in which the drug product will be marketed, including the identity of materials of construction and their compatibility with the drug product

h. Stability—describes the stability protocols and results supporting the proposed expiration-dating period and storage conditions; provides the postapproval stability protocol and stability commitment

i. The remaining information in the pharmaceutical quality section refers to samples of drug substance, finished product, and/or reference standard (if requested during the review), validation of analytical methods, and environmental impact. It is important that these subsections provide adequate detail so that FDA quality reviewers can thoroughly assess any scientific matter that may have a bearing on drug product performance.

5. *Nonclinical pharmacology and toxicology*—All animal and laboratory studies relevant to review of the NDA must be included. During review of this section, the FDA reviewers look for toxic effects that may be inconsistent or inadequately characterized.[33] Acute, subacute, and chronic toxicity, pharmacological activities, and potential carcinogenicity and teratogenicity must be addressed. The section should also include a statement that avers good laboratory practices (GLPs) for each study or a brief explanation for any instance of deviation from GLPs.
6. *Human pharmacokinetics and bioavailability*—Clinical pharmacokinetic and bioavailability data and analyses should be detailed, and relevant analytical and statistical methods must be described. A reasoned discussion of the pharmacokinetics, metabolism, and bioavailability of the drug product is essential.[34]
7. *Clinical microbiology*—A clinical microbiology section is provided only when effects on the physiology of a targeted microorganism are relevant to the review of the NDA (eg, review of an anti-infective drug product). This section addresses the drug action on microbial physiology, the antimicrobial spectrum affected, resistance mechanisms, and laboratory methods used. The possibility of sepsis arising from products should be addressed (eg, liquids that can support microbial growth or topical products that can spread bacteria).[35]
8. *Clinical data section*—There is an extensive set of FDA guidance documents related to clinical studies in NDA submissions, which cannot be summarized in detail here.[36] The clinical dataset is the basis of efficacy and safety of the proposed drug product and generally includes the following[37]:
   a. List of IND and NDA investigators supplied with the drug
   b. List of INDs and an overview of clinical studies

c. Other NDAs submitted for the same drug substance
d. Overview of the clinical studies
e. Description and analysis of clinical pharmacology studies; controlled clinical studies relevant to the proposed use of the drug product; uncontrolled studies; and other data pertaining to the evaluation of the drug product
f. Integrated summary of effectiveness (ISE), including support for the dosage and administration of the drug as proposed in the product labeling
g. Integrated summary of safety (ISS), with a discussion of adverse events, drug-drug interactions, pertinent animal data, and any other safety concerns (eg, abuse and overdose potential)
h. Integrated summary of risk-benefit (ISRB)
i. Compliance statement regarding good clinical practices (GCPs), the Institutional Review Board (IRB), and informed consent (IC)
j. Data regarding special populations, such as geriatrics, pediatrics, or patients with renal or hepatic impairment

9. *Safety update*—The first safety update report is usually filed as an amendment 4 months after the submission of the original NDA; other safety updates may be provided at other times specified by the FDA reviewing division. The format of the safety update report is similar to the ISS (see previous section) and contains safety information that arises from ongoing studies, animal studies, and other sources. Sponsors should work with the FDA to determine the content and format of these safety updates.
10. *Statistical*—This section of the NDA is closely linked to the clinical data section. Sponsors should work closely with the FDA prior to the NDA submission to come to an agreement on the format, content, tabulations, and statistical analyses to present. This section should be carefully addressed during the pre-NDA meetings with the FDA.
11–12. *Case reports*—Tabulations of cases of patient data and data elements with proper case report forms (CRFs) are essential. The detailed requirements can be discussed at the pre-NDA meetings or, if the FDA so requests, after submission of the NDA. The CRFs for patients who died during a clinical study and for patients who discontinued due to an adverse event should also be included. FDA reviewers may require additional CRFs as well.[37]
13–14. *Patent information and certification*—Patent certification is required for any relevant patents that claim the listed drug or any other drug on which the investigations relied, or that claim a use for the listed drug or any other drug.
15. *Establishment description*—This section provides general information about the organization, physical plant, and major equipment, as well as quality assurance functions.
16. *Debarment certification*—The FDA is authorized to debar individuals convicted of crimes related to the development, approval, or regulation of drugs. The individual should not provide any type of services to sponsors of an application.[38]
17. *Field copy certification*—Applicants based in the United States must submit a "field" copy of the pharmaceutical quality (CMC) section, application form, and summary of the NDA to the relevant FDA district office. The information will be used during the pre-approval inspection (PAI) at the manufacturing site. Within the NDA, the applicant must certify that an exact copy of the CMC section has been sent to the district office.
18. *User fee cover sheet*—FDA Form 3397 is used to determine the applicability of a user fee and indicate whether a check has been mailed to an FDA account concurrent with submission of the NDA.[39]
19. *Financial information*—Clinical investigator financial disclosures (FDA Form 3455) and certification (FDA Form 3454) must be included, describing investigator financial interests and sponsor-investigator financial arrangements that could affect the reliability of the submitted data.[40]
20. *Other information*—Any information that was submitted before the NDA should be referenced in this section. A sponsor may propose other uses of this section at the pre-NDA meeting.

38.2.3 The CTD format

The FDA highly recommends that sponsors submit all marketing applications in CTD format (see next). CTD format does not alter the data or information required for the NDA in the United States, as described previously.[41] The FDA also recommends (and beginning in 2017, will require) electronic submission of CTD-formatted applications (ie, "eCTD" submissions).[42]

The CTD is a product of the International Conference on Harmonization (ICH), an international organization composed of the pharmaceutical industry and regulatory representatives from the European Union, Japan, and the United States. The CTD provides a common format for submission of marketing applications submitted to the vast majority of agencies worldwide. The format

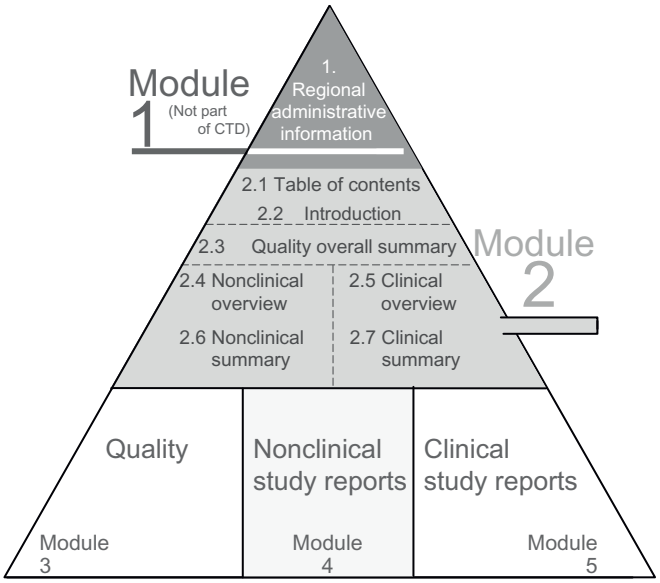


FIGURE 38.5 The common technical document (CTD) format (with Module 1 added).

and contents of CTD includes five modules, organized in the so-called "CTD triangle" diagram, as shown in Fig. 38.5.

Module 1 of CTD contains the regional administrative forms and documents as determined by each ICH region. In applications submitted to the FDA, this section contains the cover letter, form 356(h), complete table of contents of the entire CTD, and other administrative documents, such as the field copy certification, debarment certification, patent information, labeling, and the annotated labeling text. Module 2 contains the CTD summaries of information from various sections of the CTD, including a very short introduction; QOS; nonclinical overview; clinical overview; nonclinical written and tabulated summaries for pharmacology, pharmacokinetics, and toxicology; and clinical summary describing biopharmaceutics and associated analytical methods, clinical pharmacology studies, clinical efficacy, clinical safety, and a synopsis of the individual studies. Module 3 contains all of the pharmaceutical quality documents for the chemistry, manufacture, and controls of the drug substance, the drug product, and container closure system. Modules 4 and 5 contain copies of the final nonclinical and clinical study reports, respectively. The correspondence between the CTD modules and sections in an application (as described in the previous section of this chapter) is depicted in Fig. 38.6.

38.2.4 NDA review practices

38.2.4.1 Filing

Prior to submitting an NDA to the FDA, sponsors are strongly encouraged to request a meeting with the appropriate FDA review division and discuss the planned content of the application.[43] As previously discussed for the clinical phase of development (see IND Section 38.2.1), meetings between sponsors of a marketing application and FDA reviewers will help ensure that all submitted applications are complete and fileable. Marketing applications received by the FDA "central document room" are assigned a number, and the date of receipt starts the review time clock (see NDA review timelines in Fig. 38.7). When the NDA is ready for review, a regulatory project manager (RPM; note distinction from the OPQ RBPM) is charged with managing the review process and coordinating all communications with the sponsor accordingly. The RPM conducts an initial assessment (eg, whether a user fee has been paid, waived, or exempted). If all required fees have been paid and the information submitted is administratively complete, the application is then assigned to a multidisciplinary team of FDA reviewers, who by day 45 of the review (day 30 for priority reviews) must convene and discuss their respective decisions regarding the fileability of the application. All filing review issues identified by the review team are conveyed to the applicant in a single communication, which will be documented in writing and archived according to CDER standard procedures. Applications that omit necessary information not readily rectified may result in a refuse-to-file (RTF) letter that must be issued by day 60. Otherwise, a filing letter is issued and review activities are triggered within the FDA.[44]

38.2.4.2 Review

The overall NDA review and approval process (or "cycle") usually takes place in six major steps (presubmission activities, process submission, review plan, conduct review, take official action, and postaction feedback) as illustrated in Figs. 38.7 and 38.8.[45,46] Once the application is deemed fileable, a planning meeting is held with the review team to discuss timelines, high-level labeling revisions, need for advisory committees and/or inspections, and review activities. The review timelines for NDAs that are not governed under PDUFA V are 10 months from receipt for standard reviews and 6 months for priority reviews. NMEs and BLAs, as part of the PDUFA V "program," are reviewed within 12 months for standard submissions and 8 months for priority submissions from the date of submission (10 months and respectively, 6 months from the 60-day filing date). During the main review phase, each assigned reviewer, according to discipline (eg, medical officer, pharmacologist, chemist, statistician, microbiologist, clinical pharmacologist), assesses the pertinent portion of the application, proposes labeling revisions, and writes a review. Reviewers consult with each other and with their leads and communicate information requests (IR) to the applicant on a regular basis.

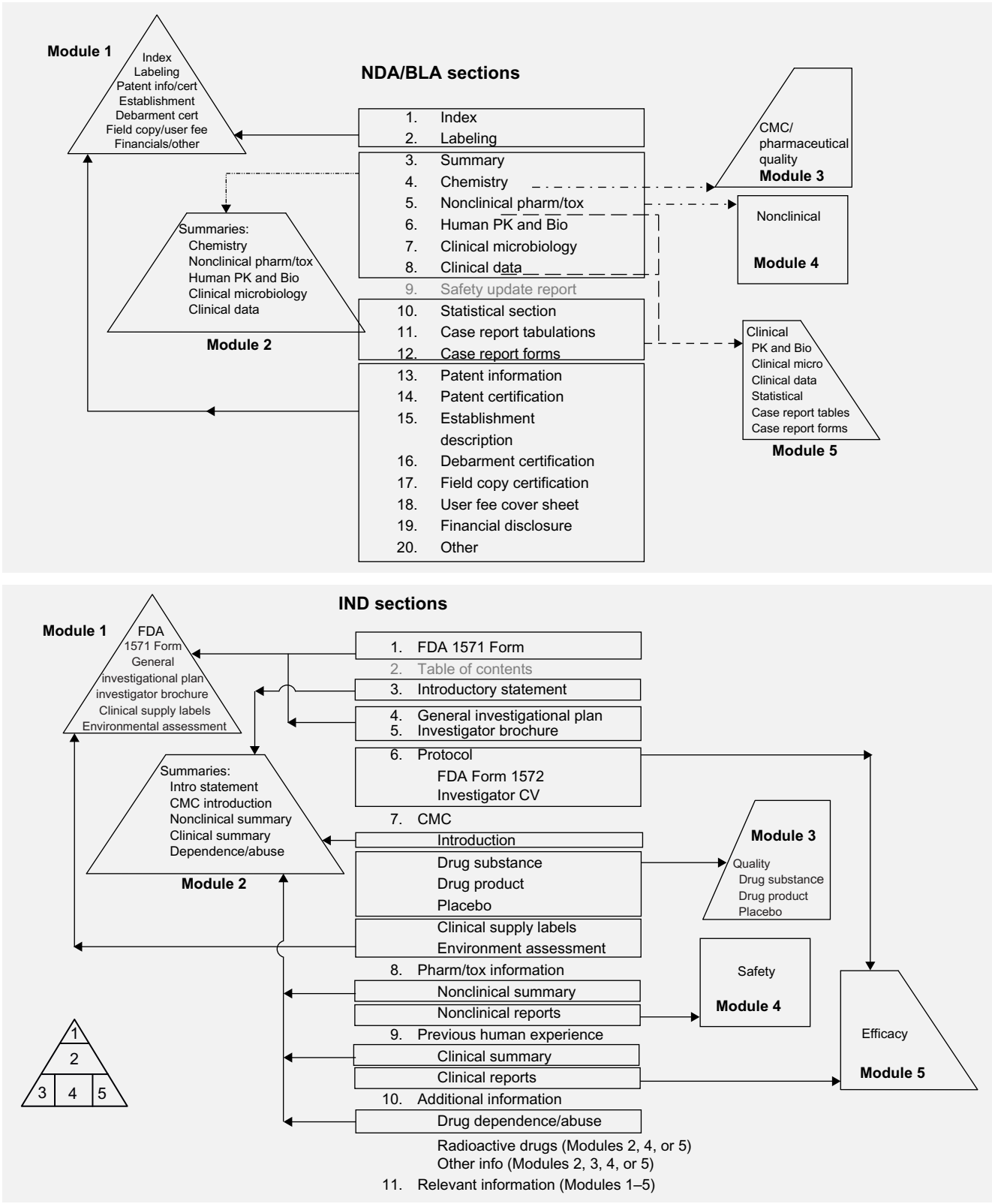


FIGURE 38.6 Mapping the NDA and IND into the five modules of the CTD triangle. *Source: Adapted from FDA internal presentation eCTD Clinic for CDER Reviewers and PMs by Valerie M. Gooding.*

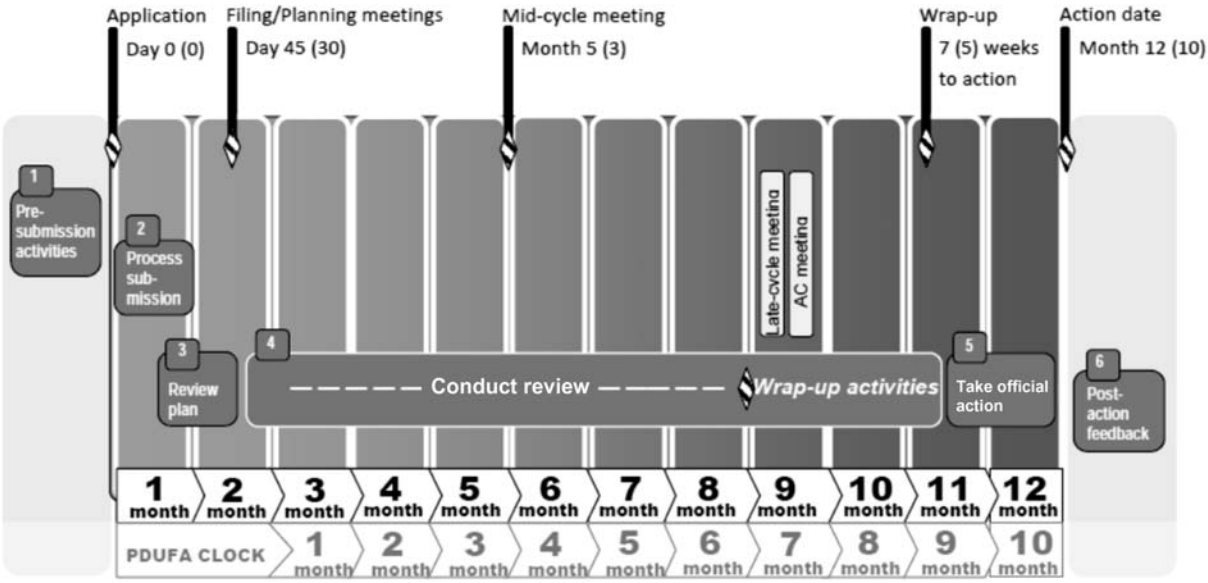


FIGURE 38.7 High-level illustration of the timeline for NDA/BLA reviews under PDUFA V. (The timeline for review of NMEs/BLAs under PDUFA V "program" extends "conduct review" phase by 2 months. Numbers in parentheses indicate priority review.)

All technical and high-level labeling issues that are critical to the approval of the application are identified by the time the mid-cycle meeting is held (ie, month 5 for standard reviews and month 3 for priority reviews). In addition to mid-cycle review, a late-cycle meeting between the review team and the applicant is generally held for standard PDUFA V "program" reviews no later than 3 months prior to the PDUFA V goal date (2 months prior to the PDUFA date for priority reviews). Deficiencies identified by respective disciplinary reviewers are usually communicated to the applicant in advance of the late-cycle meeting, although applications that involve complex issues may extend the review period by 2 months. The outcomes of all review activities, including inspection results, are integrated during the wrap-up meeting; the medical officer takes the lead to reconcile observations and written summaries, after which the action package is finalized.

In the event that the application is approved, the sponsor receives a letter that authorizes the manufacturer to distribute product in accordance with a list of agreements and any postmarketing commitments. If the agency decides on the basis of its review not to issue an approval letter, a complete response (CR) letter will be issued, citing deficiencies and offering recommendations. Mechanisms for additional feedback, after issuance of an approval letter or a CR letter, may be exercised in some cases. Postapproval meetings are offered for all NMEs (and original BLAs) and may otherwise be provided at the agency's discretion for the purpose of learning from the review experience. Applicants who receive a CR letter may request an end-of-review conference to be scheduled by the FDA to discuss deficiencies and further steps to take, including resubmission; withdrawal; or request an opportunity for hearing.[47]

38.2.4.3 Labeling

Final discussions about the package insert will usually take place near the end of the review period, possibly 2–4 weeks before the PDUFA action date (see Fig. 38.7). At this time, the FDA will communicate recommendations concerning the product label to the sponsor; and the agency and sponsor will work to reach final wording prior to the action date. The final negotiations for the labeling are often done by email or teleconference. The sponsor submits the final version as an amendment to the NDA in an electronic format.

38.2.4.4 FDA-sponsor communications during NDA review[46]

While evaluating the quality, safety, and efficacy of the drug product, based on a thorough scientific examination of the submitted application, reviewers may choose to seek clarification from the sponsor in the form of an information request email or teleconference, discipline review letter, or face-to-face meeting (appropriately documented within the agency). The RPM is involved, by phone or mail, in mediating status updates. Any information needed in support of application review must be officially submitted as an amendment to the application. Under present guidelines, communication per email is not official.

38.2.4.5 Advisory committee

If significant issues arise during the review process, the review team can seek advice from an advisory committee.

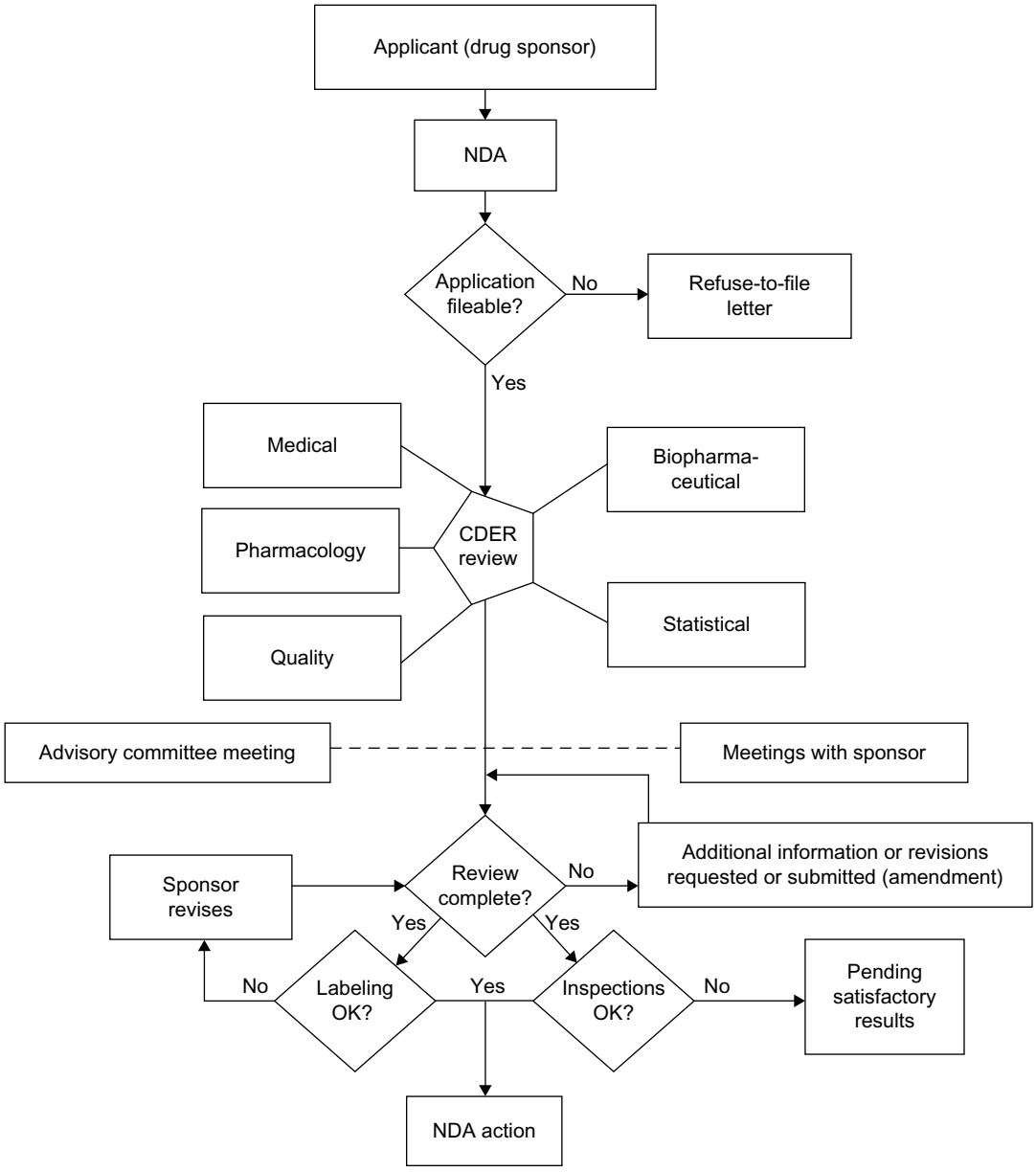


FIGURE 38.8 The NDA review process.

An advisory committee comprises a specialized group of external experts,[48] identified by the FDA, that convene to provide advice to the agency concerning any number of matters. Advisory committee meetings may be convened, for example, to discuss a NME; novel clinical trial designs or use of surrogate end points; particular safety issues or drug effectiveness in specific contexts; or public health questions arising from a proposed indication.

38.2.5 Special approval pathways for NDAs/BLAs

38.2.5.1 Expedited approval pathways

In order to increase the availability of drugs intended to treat serious and/or life-threatening conditions, four expedited drug-approval pathways have been created: priority review designation, accelerated approval, fast-track designation, and breakthrough therapy designation.[49] In 2014, among the 41 NDA/BLA applications approved by CDER, 27 (66%) fell within one of these expedited programs. Table 38.5 gives an overview of these programs. It should be noted that regulatory standards for efficacy, safety, and quality are to be rigorously upheld in all FDA review programs. Increased attention lately has been given to the breakthrough pathway, designed to expedite the development of drugs for life-threatening conditions or unmet medical needs. Scientific and logistic challenges posed by accelerated development and review can be alleviated by: (1) communication between sponsors

TABLE 38.5 Overview and Comparison of Expedited Approval Programs at FDA[49]

Program	Priority review designation	Accelerated approval approval pathway	Fast-track designation	Breakthrough therapy designation
Statuary reference	1992 PDUFA	21 CFR 314 & 601, FDCA 506(c)/ FDASIA	FDCA 506(b)/1997 FDAMA/FDASIA	FDCA 506(a)/FDASIA
Qualifying criteria	Serious condition and significant improvement in safety and effectiveness, if approved	Serious and life-threatening condition; and meaningful advantage over available therapies; effect on surrogate or intermediate endpoint reasonably likely to predict clinical benefit	Serious and life-threatening condition; and meet unmet medical needs	Serious and life-threatening condition; and substantial improvement on clinical significant endpoint over available therapies
Features	6-month review clock (compared with standard 10 month)	Approval may base on surrogate endpoint or intermediate clinical endpoint Confirmatory study obligation	Rolling review Actions to expedite development and review (eg, more frequent interactions)	Actions to expedite review Intensive FDA guidance Rolling review Organizational commitment Involvement of senior management

and the FDA regarding risk mitigation strategies to manage risk to product quality; (2) streamlining the review of stability data; and (3) using postmarketing commitments to address postapproval risk.[50]

38.2.5.2 *Orphan drug designation program*

An orphan designation signifies that a drug indication applies to a rare disease, which by definition affects fewer than 200,000 American patients per year. The Orphan Drug Act was enacted in 1983, recognizing that sponsors had little financial incentive to develop orphan drugs. The Orphan Drug Act (1983) and amendments (1992 and 2013) established several incentives to encourage the development of orphan drugs, including market exclusivity for 7 years after approval, fee waivers, tax credit on clinical research, technical assistance during the review process, as well as other federal research grants.[51]

38.2.5.3 *Pediatric exclusivity*

There have been a few FDA initiatives to improve medical product research and availability for children, including the FDA Modernization Act (FDAMA) of 1997, Best Pharmaceuticals for Children Act of 2002, FDA Amendments Act (FDAAA) of 2007, and FDASIA of 2012. Pediatric exclusivity is a 6-month extension of exclusivity added to existing periods of marketing exclusivity or patent protection.[52] The FDA is actively engaged in scientific programs that strive to improve access to drug products proven to be safe and effective for children.

38.3 THE ABBREVIATED NEW DRUG APPLICATION PROCESS

ANDAs contain information submitted to CDER for the review of generic drug products. Generic drug applications are "abbreviated" because they generally allow the sponsor to rely on preclinical (animal, cell culture) and clinical (human) data used to establish the safety and effectiveness of a RLD. In this way, generic products must be adjudicated as therapeutically equivalent to an approved drug product (ie, the RLD). Therapeutic equivalence must be established through the combined criteria of pharmaceutic equivalence and bioequivalence. Pharmaceutic equivalence signifies that the generic product and the RLD contain the identical API, API strength, dosage form, quality, performance characteristics, and that the route of administration intended use are the same for the generic and RLD.[53] Bioequivalence signifies that the generic product exposes the patient to the same amount of API in the same time course as the RLD, observed on the basis of plasma drug levels (see Section 38.3.3.2). Generic applications are filed under FDCA section 505(j) and authorized for marketing only after all forms of exclusivity for the innovator drug product have expired.

38.3.1 Format and content of the ANDA

The format and content of ANDA submissions to the FDA are specified in regulations (21CFR 314.94) and reflected in the "harmonized" application form, FDA 356(h). Table 38.6 gives the list of contents of an ANDA (described next) and shows the difference between the required contents of NDA and ANDA submissions.

1. *Index/table of contents*
2. *Basis for ANDA submission*—shows that the ANDA refers to the RLD selected by the FDA for conducting bioequivalence testing and documents a comparison between the generic drug and the RLD, including a statement according to which conditions of use recommended in the labeling proposed for the generic drug product have been previously

TABLE 38.6 Comparison of NDA and ANDA Contents

NDA	ANDA
	Basis of submission
	Comparison between generic and RLD
	Request for waiver for in vivo BA/BE studies (if applicable)
Labeling	Labeling
Summary	
Chemistry	Chemistry
Nonclinical pharmacology and toxicology	
Human pharmacokinetics and bioavailability	Human pharmacokinetics and bioavailability/bioequivalence
Clinical microbiology (if applicable)	
Clinical data	
Safety update report	
Statistical section	
Case report tabulations	
Case report forms	Case report form
Patent information	Patent information
Patent certification	Patent certification
Establishment description	
Debarment certification	Debarment certification
Field copy certification	Field copy certification
User fee cover sheet	
Financial information	Financial information
Other	References

approved for the RLD; states that the active ingredient is the same as that used in the RLD and that the route of administration, dosage form, and strength of the proposed drug product are the same as those of the RLD.

3. *Labeling*—according to 21CFR314.94, includes a side-by-side comparison of the applicant's proposed labeling (container and carton) and labeling approved for the RLD for each strength and package size; any differences (eg, those filed by petition under section 314.93 or differences owing to exclusivity under patent or under section 505(j)(4)(D) of the Act) should be explained in annotation. All labeling for a generic drug product must comply with the PLR and be provided in SPL formatting.
4. *CMC/pharmaceutical quality*—contains information as required under 21CFR 314.50(d).

The quality section of the ANDA, like that of the NDA, covers raw materials (eg, API, reagents, and excipients), manufacturing facility, manufacturing of the API and in-process controls, materials used in API packaging, controls for the finished dosage form, analytical methods, stability of the finished dosage form, sterilization assurance data, as well as samples and environmental considerations, generic enforcement act and a US agent letter of authorization (LOA). Stability requirements for generics and new drugs are identical, per ICH and FDA guidelines. A master production record (proposed or actual) and a description of the equipment to be used in the manufacture of the commercial batch of the drug product should also be provided.

There are differences between the pharmaceutical quality/CMC review for an ANDA and that for an NDA. The ANDA sponsor must typically reference a DMF previously submitted to the FDA containing proprietary information because the drug substances used in the manufacture of drug products that are the subject of ANDAs are usually obtained from third-party external manufacturers (see Section 38.3.4.2). There are also additional requirements for ANDAs that do not necessarily appear in NDAs:

- **a.** Any inactive ingredient in the proposed generic product that differs from the RLD should be identified and demonstrated as having no effect on safety and efficacy
- **b.** Specifications should be justified in reference to compendia (eg, USP, JP), ICH, and/or analysis of the RLD
- **c.** Release characteristics of the proposed generic (ie, dissolution) should be comparable to those of the RLD

Notably, in order to modernize the science- and risk-based pharmaceutical quality assessment of ANDA applications, the FDA has implemented a QbR approach (see Section 38.1.5.2), and sponsors are now generally expected to submit a QOS employing the QbR format.[54] QbR assists applicants in developing the QOS by providing specific questions to ensure that adequate information associated with the robust manufacture and reliable control is submitted for FDA review.

5. *Human pharmacokinetics and bioavailability/bioequivalence*—contains information that shows that the drug product is bioequivalent to the RLD.

According to current guidance from CDER, conclusions that two drug products are bioequivalent should reflect significant agreement in pharmacokinetic parameters (ie, area under the curve (AUC) and maximum concentration (C_{max}))

such that the entire 90% confidence interval associated with the generic-to-reference ratio of geometric means should fall within the bioequivalence limits of 80% to 125%. Exceptions exist, such as in the case of highly variable drugs and those with narrow therapeutic indices, and waivers of in vivo bioequivalence studies are also available in certain cases[55,56] (see discussions on bioequivalence in chapter: "Bioavailability and Bioequivalence" of this edition). Waivers of clinical bioequivalence studies (ie, biowaivers) can also be granted in some cases (eg, for parenteral solution or ophthalmic solution that is qualitatively (Q1) and quantitatively (Q2) the same as the RLD, or for highly soluble, or highly permeable, or rapidly dissolving drug products based on biopharmaceutics classification system (BCS)).[57]

6. *Case report forms*—(see case reports for NDAs)
7. *Patent information and certification*—contains all relevant patents for drug product, method of use, method of manufacturing, licensing agreements, disputed patent information, amended certifications, etc.

   ANDA sponsors must confirm patent certification under one of the following criteria (ie, paragraphs):
   - **a.** That no patent information on the drug product that is the subject of the ANDA has been submitted to the FDA;
   - **b.** That any such patent has expired;
   - **c.** The date on which such patent expires; or
   - **d.** That such patent is invalid or will not be infringed by the manufacture, use, or sale of the drug product from which the ANDA is submitted.
8. *Financial certification or disclosure statement* (see certification and financial information for NDAs).

   An applicant may amend an ANDA under review to revise existing information or respond to IR. In general, IR tend to involve the CMC (stability, specifications, analytical methods) or BE sections and are the major cause of delays in ANDA approvals. Applicants should take care to complete these sections prior to submission.

38.3.2 The CTD format of an ANDA

The FDA strongly recommends that all ANDAs be electronically submitted in the CTD format (eCTD); submission in eCTD format will be required as of May 15, 2017. ANDA submissions that are not in eCTD format are excluded from current GDUFA guidelines governing timelines for FDA review. The CTD format for ANDAs is similar to that for NDAs; however, the order of presentation (detailed next)[58] differs slightly.

Module 1 of the ANDA CTD submission contains several forms (ie, 356(h) (see previous NDA discussion), the GDUFA user fee cover sheet (FDA Form 3794), and Form FDA 3674), a cover letter, and administrative documents related to US agent, field copy certification, debarment certification, financial certifications, patent information and certification, as well as statements of right of reference and LOA for each DMF referenced in the ANDA. In addition, Module 1 contains information demonstrating that the generic product is the same as the RLD and if applicable, any biowaiver request. Labeling information as previously described is also submitted in Module 1. Module 2 contains the QOS, which provides an overview of the CMC section as related to drug substance and drug product; summary reports for bioequivalence studies are also submitted in Module 2. Similar to NDA submissions, Module 3 for ANDAs contains all the CMC information in the application. For drug substance, the applicant should provide general information, characterization, control of drug substance (eg, specifications, descriptions of analytical procedures, and validations), batch analysis with certificate of analysis (COA), information about reference standards or materials, container closure systems, and stability data, including retest period or expiration date. For drug product, the applicant should provide description and composition, information on the pharmaceutical development (including the pharmaceutical development report and microbial attributes), information about the manufacture (including manufacturing process and controls), information on the controls of excipients (including source and grades), information supporting the controls for the release of the drug product, information on reference standards or materials and container closure system, as well as stability data, postapproval stability protocols and stability commitments. Module 4 data do not generally apply to ANDAs. Module 5 contains the clinical study report data needed to demonstrate bioequivalence to the RLD.

38.3.3 ANDA review practices

38.3.3.1 Filing

Prior to submission, ANDA sponsors may choose to consult the FDA in the form of a controlled correspondence by requesting information on a specific element of generic drug product development, especially for first-generic products.[59] Formal meetings between the ANDA sponsors and the FDA before submission are not as common as for NDA sponsors, but they may be granted in special cases. Upon submission, the FDA determines whether each ANDA is sufficiently

complete to permit a substantive review. If the submission is incomplete, the agency can issue a refuse-to-receive (RTR) letter, which avoids extended review period, multiple review cycles and waste of resources. Common deficiencies that will elicit a RTR letter include: failure to provide the 356(h) form; nonpayment of user fees; lack of a designated US agent (for a foreign applicant); failure to provide the environmental assessment or claim of categorical exclusion; citing a suitability petition (ie, noting changes from the RLD) that is pending; lack of information for API and/or drug product (eg, inadequate stability data, batch records, or methods validation); and bioequivalence issues (eg, inadequate data, biowaiver, or dissolution). Complete ANDA submissions may be filed under Paragraph I, II, III, and IV certifications as listed in the "orange book" for the RLD (see Patent Information). An ANDA applicant who files a Paragraph IV certification must, within 20 days of filing, notify the innovator, who will have 45 days to take action upon receiving the notification. The FDA may hold the application up to 30 months, depending upon the outcome of the litigation between the innovator and the ANDA applicant, if any. A regulatory filing checklist for ANDA submissions is regularly updated on the FDA website.[60]

38.3.3.2 *Review*

As more and more RLDs have come off patent, and as more companies have entered generic drug manufacturing, the volume of ANDAs has been increasing over the past decade, resulting in a considerable backlog of submissions. An important objective of GDUFA was therefore to set timelines for review and resolve the ANDA backlog by 2017. The FDA is committed to review and act, within 15 months from submission, on 60% of those original ANDAs submitted between Oct. 1, 2014 and Sep. 30, 2015 (ie, the year 3 cohort); similarly, 75% of those submitted between Oct. 1, 2015 and Sep. 30, 2016 (ie, the year 4 cohort), must receive action. The agency must act within 10 months on 90% of submissions made between Oct. 1, 2016 and Sep. 30, 2017 (ie, the year 5 cohort). For year 1 and year 2 cohorts (ie, received between Oct. 1, 2012 and Sep. 30, 2013 or between Oct. 1, 2013 and Sep. 30, 2014, respectively), the FDA will expedite review of Paragraph IV applications.[61] GDUFA review performance goals are depicted in Table 38.7, and ANDA review timelines for cohort 5 are schematized in Fig. 38.9. To ameliorate ANDA burdens and assist ANDA applicants in improving the quality of submissions, the FDA has issued a number of guidance documents.[62]

The bioequivalence, labeling, and CMC/quality sections of the ANDA are the major components of review. Bioequivalence is evaluated by comparing the extent of absorption of test and the reference products, in crossover clinical studies on healthy subjects, as determined by the AUC of plasma concentrations plotted over time. Both the AUC and peak plasma concentration C_{max} are used in the determination of bioequivalence. Recommendations of study designs and data evaluation for bioequivalence study, as well as appropriate in vitro dissolution methods, are listed in the FDA regulatory guidance.[64] Review of labeling text focuses on differences in excipients, presentation of pharmacokinetic data, clarity of information, and the utility of labeling language with respect to pharmacy practice. Reviewers of the CMC/quality information, organized within a number of disciplines, work in parallel to assess the quality of the drug substance and drug product, manufacturing and controls, batch formulation and records, facilities, product specifications, packaging, and stability. A number of review initiatives are underway

TABLE 38.7 GDUFA Review Performance Goals[63]

	FY 2013	FY 2014	FY 2015	FY 2016	FY 2017
Original ANDA	Expedite review of paragraph IV and maintain pre-GDUFA productivity		60% in 15 months	75% in 15 months	90% in 10 months
Tier 1 first major amendment	Maintain pre-GDUFA productivity		60% in 10 months	75% in 10 months	90% in 10 months
Tier 1 minor amendments (1st—3rd)	Maintain pre-GDUFA productivity		60% in 3 months[a]	75% in 3 months[a]	90% in 3 months[a]
Tier 1 minor amendments (4th—5th)	Maintain pre-GDUFA productivity		60% in 6 months[a]	75% in 6 months[a]	90% in 6 months[a]
Tier 2 amendment	Maintain pre-GDUFA productivity		60% in 12 months	75% in 12 months	90% in 12 months
Prior-approval supplements	Maintain pre-GDUFA productivity		60% in 6 months[a]	75% in 6 months[a]	90% in 6 months[a]
ANDA amendment and PAS in backlog on Oct. 1, 2012	Act on 90% by end of FY 2017				
Controlled correspondences	Maintain pre-GDUFA productivity		70% in 4 months	75% in 2 months	90% in 10 months

[a]*10 months if inspection is required.*

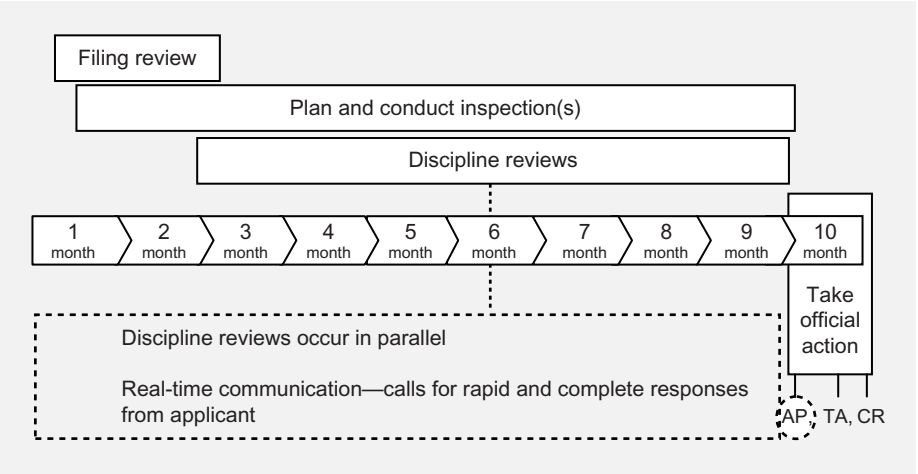


FIGURE 38.9 High-level illustration of the timeline for review of ANDA year 5 cohort.[65]

to ensure the timely review of ANDAs; some of these focus on information requests (IR), real-time communication (RTC), easily correctable deficiencies (ECD), and complete response (CR) letters. These initiatives are aimed at reducing the ANDA backlog and increasing the rate of ANDA approval.

If all review disciplines find the ANDA acceptable and if all facilities are in satisfactory standing as reviewed and inspected, the ANDA receives approval. If the ANDA review process is completed prior to the expiration of innovator drug product exclusivity, a notice of tentative approval (TA) may be issued, which does not allow the generic drug manufacturer to distribute product. An approval letter is contingent upon the resolution of litigation and issues related to product exclusivity.

38.3.4 Special considerations for ANDAs

38.3.4.1 ANDA filing and market exclusivity

New drug product exclusivity is provided by FDCA section 505(c)(3)(E) and 505(j)(5)(F). A 5-year period of exclusivity is granted to NDAs seeking to market a new chemical entity, either alone or in combination, but after 4 years, information pertaining to safety and efficacy from such an NDA approval may be used for the submission of a 505(b)(2) application or ANDA, provided that the submission supplies evidence of no infringement. If a follow-on drug product containing a previously approved active moiety is approved on the basis of new clinical investigations (other than bioavailability studies), then the follow-on product may be granted a 3-year period of exclusivity. Applicants seeking to market the first generic version of a drug product may be eligible for a 180-day period of exclusivity. Orphan drug exclusivity extends 7 years, and pediatric exclusivity gives 180 days of extension to existing patents or exclusivity.

38.3.4.2 Drug master files (DMFs)[66,67]

DMF is a document, prepared and submitted by a manufacturer that provides detailed proprietary information about facilities, process, or articles used in manufacturing processing, packaging, and storing. The submission of a DMF is not a regulatory requirement and is not a substitute for any marketing application; rather, it provides information to support an IND, NDA, ANDA, another DMF, an Export Application, or amendments and supplements to any of these. A DMF is typically filed when two or more firms work in partnership on developing or manufacturing a drug product; the DMF protects a firm's intellectual property from its partner while complying with regulatory requirements for disclosure of processing details. In the United States, DMFs can fall into one of five types[68]:

- Type I: Manufacturing site, facilities, operating procedures, and personnel (no longer applicable, such information for sterile and biotech products may be filed as a DMF Type V)
- Type II: Drug substance, drug substance intermediate, and material used in their preparation, or drug product
- Type III: Packaging material
- Type IV: Excipient, colorant, flavor, essence, or material used in their preparation
- Type V: FDA-accepted reference information

Before the FDA can review a DMF or DMF information in support of an application, the DMF holder must submit a LOA that allows the FDA to reference the DMF (the agency will review a DMF only if information in that DMF is referenced by another application). If the FDA reviewers find deficiencies in the information provided in a DMF, a letter describing the deficiencies is sent to the DMF holder. Deficiencies in a DMF cannot be disclosed to the sponsor who

references the DMF because the information is proprietary; however, the referencing sponsor is notified by the FDA if the DMF is deficient, and it is then the DMF holder's responsibility to notify the sponsor when deficiencies in the DMF have been addressed. A DMF itself is never "approved"; instead it is reviewed to determine whether it is adequate to support a particular application that references it.

DMFs are commonly used to support the drug substance information in ANDAs; such (type II) DMFs are subject to GDUFA fees, an initial "completeness assessment" (CA), and communications with DMF holders.[69] The CA review determines whether the DMF is ready to undergo a comprehensive scientific review, but does not guarantee that the DMF will be found adequate. Type II DMFs must cover characterization of the drug substance, manufacturing and controls of the drug intermediate or substance (including a list of the organic impurities, elemental impurities, and residual solvents), reference standards or materials, container closure system used for packaging the drug substance, and stability data and storage conditions. Only following the review of this information can the DMF be found adequate, allowing it to be referenced in an ANDA. A list of DMFs that have passed the completeness assessment and are available for reference by ANDAs under GDUFA is available on the FDA website.[70]

38.4 THE BIOLOGIC LICENSE APPLICATION PROCESS

Owing to the history of biological therapeutics in the United States—specifically, early vaccines and serum-derived agents—the FDA review of biologics is authorized not solely through the FDCA, but also through PHS Act section 351. Today, the spectrum of biologic drugs that are reviewed by CDER (in either the form of a BLA or an NDA) includes monoclonal antibodies and other proteins, such as cytokines, enzymes, growth factors, interferons, and nonvaccine therapeutic immunomodulators. The review of other biological products is the responsibility of the FDA CBER and includes: vaccines, blood and blood-derived components or products, plasma products, human tissues and cellular products, allergen patch tests, allergenic extracts, antitoxins, in vivo diagnostics, toxoids intended for immunization, and gene therapy.[71]

38.4.1 Format and content of the BLA

The harmonized application form FDA 356(h) used for all drug products (see previous sections of this chapter) also specifies format and content requirements for a BLA for a biologic product. The sections expected in the BLA are thus identical to those in the NDA: index, summary, CMC/pharmaceutical quality, establishment description, nonclinical pharmacology and toxicology, human pharmacokinetics and bioavailability, clinical microbiology, clinical data, statistical data, case report forms and tabulations, labeling, patent information, and establishment description. The CMC/pharmaceutical quality section in a BLA must present CMC information for both the drug substance and the finished dosage form of the product. Regardless of whether production is by fermentation, cultivation, isolation, or synthesis, the CMC information for drug substance includes description and characterization, manufacture and process control, manufacturing consistency, drug substance specifications, reprocessing, packaging, and stability. With regard to the drug product, the CMC section of the BLA must include information on composition and characterization, manufacturer and facilities, manufacturing methods, drug product specifications, container closure system, microbiology, lyophilization, and drug product stability. Typically, CMC subsection of the BLA is more detailed for the drug substance than for the drug product. Biologic drug products dosage forms are typically liquid or lyophilized, and the manufacturing process has relatively few steps. In contrast, drug substance manufacturing involves multiple steps needed to purify the intended product from a complex matrix composed of cells and other proteins. For an assessment of biosimilarity, BLA requirements additionally include data derived from analytical studies, animal studies (including assessment of toxicity), and clinical studies (including assessments of immunogenicity and pharmacokinetics/pharmacodynamics).[72] For details on the content of each section in an application, see Section 38.2.2 of this chapter.

38.4.2 BLA review practices

Biologic drugs conform to the FDCA definition of "drug" as given and are therefore subject to regulations that govern drug review and approval. Clinical studies of biologics in humans may be initiated and pursued under an IND in accordance with 21 CFR 312 (see Section 38.2.1), the results of which, along with relevant nonclinical data, can be submitted as part of a BLA. In reviewing a BLA, the FDA assesses whether the product, the manufacturing process, and the manufacturing facilities meet applicable requirements to ensure the continued safety, purity, potency, and quality of the product; the PHS Act emphasizes the importance of appropriate manufacturing controls for products, particularly where changes to the biological molecule that arise during manufacturing might not be detected by standard characterization techniques, and additional clinical studies are needed.

Similar to NDA review, the BLA review process is divided into six phases: (1) presubmission activities; (2) processing of submission; (3) plan of review; (4) conducting the review; (5) official action; and (6) post-action feedback (descriptions of these phases are provided in Section 38.2.4. of this chapter).[46] After receipt, the FDA must take action on submitted BLAs, for which a PDUFA V program is not specified, by 6 months (in the case of a priority review) or 10 months (for a standard review). (The timelines for NMEs and BLAs that fall under the PDUFA V program are discussed in Section 38.2.4.)

38.4.3 Special considerations for BLAs

38.4.3.1 Biosimilars

Section 351(k) of the PHS Act creates an abbreviated licensure pathway for biological products that are demonstrated to be "biosimilar" to or "interchangeable" with a biological product licensed under section 351(a) of the PHS Act.[73] Similar to the generic drug product approval, this pathway encourages competition and promotes affordable biologic drug products. Under the BPCI Act, a biologic is "biosimilar" if it is "highly similar to the reference product notwithstanding minor differences in clinically inactive components" and shows that "there are no clinically meaningful differences between the biological product and the reference product in terms of the safety, purity, and potency of the product."[72] The FDA applies rigorous scientific standards to assess biosimilarity and to make a determination that the biosimilar product is highly similar to the reference product and that there are no clinically meaningful differences between the biosimilar and the reference product.

Despite the analogy between biosimilar and generic products, there are fundamental differences between them.[74] Whereas the drug substances in generic drugs generally are homogeneous small molecules (of a single, readily defined structure), biologic drugs comprise large molecules, often glycosylated and subjected to posttranslational modifications. Consequently, a large number of analytical methods, including orthogonal assays, are necessary to characterize biologic drugs. In a series of recent guidelines on biosimilars,[73,75–78] the FDA has presented important scientific considerations for demonstrating biosimilarity, describing a stepwise approach that includes comparative studies of the proposed biosimilar and the reference product that evaluates structure, function, animal toxicity, human pharmacokinetic and pharmacodynamic data, clinical immunogenicity, and clinical data for safety and effectiveness. In the stepwise approach, the analytical, structural, and functional studies are the foundational elements in determining biosimilarity. These studies include comparative side-by-side analysis of primary, secondary, and higher-order structure, purity and impurity of the product, posttranslational modifications, both natural (eg, glycosylation) or chemical (eg, PEGylation), as well as an evaluation of the biological function and mechanism of action. At each step, sponsors should evaluate the results of the studies, identify any residual uncertainty, and design additional studies (nonclinical and clinical) to address the residual uncertainty.

The FDA has established biosimilar product development (BPD) meetings for sponsors to seek advice from the agency on certain aspects of the development program (analytical, nonclinical, and clinical) and obtain preliminary feedback on the similarity of the proposed product to the reference product. To ensure a successful development program, the FDA recommends that sponsors meet with the agency in early stages and at milestone steps throughout product development.[72] At the completion of the development program, sponsors submit a 351(k) BLA. The review process is the same for a 351(a) and a 351(k) BLA. The FDA will make a final determination that the proposed product is biosimilar to the reference product based on the totality of the evidence provided in the 351(k) submission.

38.4.3.2 Biosimilars and market exclusivity

Under the BPCI Act, a given marketed biologic (ie, the reference product) is granted a period of exclusivity in which the FDA cannot approve a biosimilar application (351(k) BLA) to a reference product for a period of 12 years. Furthermore, the FDA cannot file a 351(k) BLA earlier than 4 years from the expiration of the exclusivity period for the reference product. The Purple Book, published by FDA, lists licensed biological products, and their corresponding periods of exclusivity.[79]

38.4.3.3 BPCI act and protein products under NDA

The statutory definition of "biological product" evolved over the years. It was initially defined in the Biologic Control Act of 1902, and subsequently modified by the legislature in the PHS Act of 1944 to read "…any virus, therapeutic serum, toxin, antitoxin or analogous product applicable to the prevention and cure of diseases of man." The Food and Drug Act of 1906 and the Food, Drug, and Cosmetics Act of 1938 defined "drug" broadly to include, among other things, substances intended for use in the cure, mitigation, or prevention of diseases. Protein products could fall under either "biologic" or "drug" definition, and the FDA assigned jurisdiction to centers case-by-case until 1991, when an intercenter agreement clarified product jurisdiction between the Center for Drugs and

the Center for Biologics, with products assigned to CDER regulated under NDA and products assigned to CBER regulated under BLAs. As a result, some protein products were assigned to CDER and were regulated under NDA. With the BPCI Act of 2009, Congress further modified the definition of biological product, to include "protein, except any chemically synthesized polypeptide." With the new statutory definition of biological product, certain protein products, approved under section 505 of the FDCA, are now considered biological products. Congress ruled that within 10 years enactment of the BPCI, the protein products approved under NDA would be deemed BLAs. As of Mar. 23, 2020, marketing applications for a biological product (including forms of insulin, hyaluronidase, menotropins, and human growth hormones) that are currently approved as drugs under section 505 of the FDCA, will be regulated as biologics under section 351 of the PHS Act.[76]

38.5 POSTAPPROVAL ACTIVITIES AND LIFE CYCLE MANAGEMENT OF NDAs, ANDAs, AND BLAs

Whereas "amendments" are submitted to update or modify unapproved applications (IND, NDA, BLA, ANDA), "supplements" are submitted to modify approved/licensed applications. Supplements must address postapproval changes to components and composition, manufacturing sites, manufacturing processes, specifications, container closure systems, labeling, and other various aspects (eg, stability protocols or expiration dates). Supplement submissions must list all changes that may affect the drug product, with sufficient detail for the FDA to readily determine the scope of change. As described in 21 CFR 314.70, postapproval changes to an approved NDA, ANDA, or BLA are classified as major, moderate, or minor, and should be reported using one of four reporting categories[80]: prior-approval supplements (PAS), CBE-30 and CBE-0 (or CBE-immediate) supplements, or annual reports. In a recently released draft guidance, the FDA clarified which elements of the CMC information in an application constitute "established conditions" that, if changed following approval, require reporting to FDA, and where in the application these elements are generally expected to be described.[81] Established conditions include description of product, manufacturing process, facilities and equipment, and elements of the associated control strategy, as defined in the approved application, that assure process performance and product quality; any changes in established conditions must be reported as per 21 CFR 314.70 and 601.12. When finalized, this guidance will allow for a more effective postapproval submission strategy, including risk-management principles and knowledge management, by the sponsors as well as the FDA. The intent is to enhance reliance on sponsor's internal quality systems to properly asses, validate, and implement postapproval changes, which will eventually reduce and even eliminate certain reporting requirements.

38.5.1 Prior-approval Supplements

PAS describe major changes, having the substantial potential to adversely affect product safety, efficacy, and/or quality; such changes require approval by the FDA before their implementation is manifested in any distributed product. Postapproval changes to be described in a PAS include changes in formulation (eg, dosage or ingredient-release form), components and composition, moves to a new manufacturing site or establishment that could warrant an FDA inspection, or placement of aseptic sterilization or manufacturing processes, including packaging and closure systems, within new locations (eg, refurbishing or use of a location where similar processes have not been used in such manufacturing). Changes in other processes that must be described in a PAS include those that would affect dose delivery to the patient (eg, metering) and any change that can affect the impurity profile or the physical, chemical, or biological properties of drug product. Changes that could affect specifications or analytic processes used in assessing safety-, efficacy- or quality-related specifications must also be described in a PAS. Finally, the PAS should be used to describe any changes to product labeling (eg, based on postmarketing study results).

38.5.2 Changes being effected (CBE-30 and CBE-0 supplements)

CBE-30 supplements (for changes being effected in 30 days) describe changes having moderate potential to adversely affect safety and effectiveness; if, within 30 days of receiving a CBE-30 supplement, the FDA informs the applicant that a PAS is required, no drug product affected by the change may be distributed. CBE-0 supplements (without the 30-day delay prior to implementation) may also be used to describe moderate changes. Moderate changes to be reported in the CBE-30/CBE-0 supplements may pertain to manufacturing facilities or packaging sites for drug substance or drug product, manufacturing processes, equipment, starting materials, specifications, analytical procedures, container closure systems, or labeling. Examples can be found in the FDA's published guidelines.[80]

38.5.3 Annual reports

Annual reports may function as supplements to report minor changes, having little potential to adversely affect drug product. Changes reported in annual reports may be related to new manufacture sites for intermediates, new packaging sites for secondary packaging, new labeling sites, new equipment of the same design and operating principles, new specifications made to comply with an official compendium, tightening of specifications, container closure systems for nonsterile drug products, and minor labeling edits. Examples of each type of change that is appropriate for reporting in an annual report can be found in the FDA's published guidelines.[81,82]

38.5.4 Supplements to BLAs

BLA supplements are classified as major, moderate, and minor, depending on the potential effect on the identity, strength, quality, purity, or potency of the product; BLA supplements are reported according to the same categories listed previously for NDAs: PAS, CBE-30, CBE-0, and annual reports.[83,84] Changes that may require a PAS include, but are not limited to, changes to the sequence of the process, large scale-up of cell culture or purification, or changes to manufacturing sites. Other changes that would require submission of a supplement to a BLA can be inferred from the description of NDA supplements.

38.6 GLOBAL PERSPECTIVES ON PRODUCT REGISTRATION AND DRUG APPROVAL

Economic globalization has been especially challenging for the pharmaceutical sector, where harmonization must extend to standards and regulations involving a wide range of complex products. Over the past few decades, regulatory agencies and industrial stakeholders around the world have collaborated to standardize submission requirements and drug approval practices. Here, we briefly discuss initiatives from the ICH and regulatory bodies, such as the European Medicines Agency (EMA).

38.6.1 ICH harmonization in drug marketing submissions: the CTD format

Efforts that began in the 1980s to streamline the regulatory oversight of drug development around the world gave rise in 1990 to the ICH.

ICH includes participants from the world's three largest pharmaceutical and biotechnology markets: Europe, Japan, and the United States. The regulatory authorities involved in the ICH include the FDA and its counterparts in Europe, the EMA and in Japan, the Japanese Ministry of Health, Labor, and Welfare (MHLW). Industry representatives to the ICH include the Pharmaceutical Research and Manufacturers of America (PhRMA), the European Federation of Pharmaceutical Industries and Associations (EFPIA), and the Japan Pharmaceutical Manufacturers Association (JPMA). Observers to the ICH include representatives from the World Health Organization, Health Canada, and the European Free Trade Association (EFTA; EFTA member states are Iceland, Liechtenstein, Norway, and Switzerland).

The ICH objective is to ensure that good-quality, safe, and effective medicines are developed and registered in the most cost-effective and efficient way.[85] This objective is accomplished through the development and implementation of harmonized guidelines and standards in five steps:[86,87]

1. Building scientific consensus
2. Agreeing on draft text
3. Consulting with regional regulatory agencies
4. Adopting harmonized guidelines
5. Implementing guidelines in ICH regions

More than 50 guidelines on technical requirements have materialized from the ICH, relating to quality, safety, and efficacy, electronic standards for the transfer of regulatory information, CTD and eCTD, maintenance of ICH-controlled terminology, the medical dictionary for adverse event reporting and coding of clinical trial data (MeDRA), and efforts related to generic and OTC drugs. ICH harmonization activities extend from new guideline topics under development to the revision of existing topics needing clarification for implementation.[88] ICH topics related to the development and registration of drug products are grouped into four categories[89]:

- Quality—focuses on technical requirements for development, manufacturing, and control of the drug product, including stability testing, analytical validation, impurity profiles in drug substance and drug product, quality of biotechnology products, specifications for drug substance and drug product, GMP, pharmaceutical development, quality risk management, pharmaceutical quality system, and development and manufacture of drug substances (chemical entities and biotechnological/biological entities)[90]
- Safety—focuses on preclinical testing, in vitro and in vivo animal testing, including toxicity testing, carcinogenicity studies, genotoxicity studies,

reproductive toxicology, toxicokinetics and pharmacokinetics, pharmacology studies, immunotoxicity studies, and photosafety evaluation
- Efficacy—focuses on clinical testing, including collection and reporting of clinical data, dose-response studies, pharmacovigilance, ethnic considerations in conducting foreign trials, GCP, and studies in special populations, such as geriatrics and pediatrics
- Multidisciplinary—focuses on genotoxic impurities and nonclinical safety studies

The technical requirements outlined in ICH guidelines are discussed and agreed upon by all parties involved in harmonization before they are formally implemented by participating regulatory agencies (ie, FDA, EMA, and MHLW).

Although the use of the ICH guidelines initially targeted the ICH regions, globalization in drug development caused a growing interest in ICH guidelines and operations beyond these regions. Consequently, ICH Global Cooperation Group (GCG) is taking important steps toward establishing partnerships beyond ICH regions with involvement from Regional Harmonization Initiatives (RHIs) to promote better understanding of ICH guidelines.[91]

Finalization of the CTD (Fig. 38.5) in 2000, was a milestone accomplishment of the ICH. The CTD has alleviated costs and delays associated with converting registration files between the different national formats and removed redundancies in data sets.[92] Notably, the content of the CTD is not completely harmonized, as national or regional differences still exist in Module 1 and elsewhere throughout the submission, depending on differences in regulatory practice and procedures. For details on the CTD format and content, see Section 38.2.3 of this chapter.

38.6.2 EMA and comparison to FDA

The EMA is a decentralized body of the European Union, protecting and promoting public and animal health. The mission of the EMA is to foster scientific excellence in the evaluation and supervision of medicines, by the means of a centralized procedure that ensures coordination of EU scientific committees, inspection activities, safety monitoring of medicines (pharmacovigilance system), and referral or arbitration procedures for drugs approved via noncentralized authorization pathways.[93]

Both the FDA and EMA employ accelerated approval pathways to shorten the time needed for an application to be approved. However, whereas the FDA is authorized to issue marketing approvals on the basis of its scientific reviews, the EMA scope is limited to scientific evaluation only. Data are evaluated by the EMA advisory Committee for Medicinal Products for Human Use (CHMP), which includes independent experts nominated by each member state, as well as representatives of patient, consumer, and healthcare professional organizations. Once the scientific evaluation is completed in a certain timeline, the committee forms an opinion as to whether the drug product should be marketed, which is then transmitted to the European Commission (EC)—the authority that grants marketing authorizations in the European Union.[94] In Europe, the marketing authorization is done by the following procedures: centralized, national, decentralized, or mutual recognition.

A marketing authorization granted under the centralized procedure allows a drug product to be marketed in all the EU countries and European Economic Area (EEA) member states including Iceland, Liechtenstein, and Norway. The centralized procedure is a "reserved" procedure, not open to all products, but dedicated to human medicines for the treatment of human immunodeficiency virus (HIV) or acquired immune deficiency syndrome (AIDS), cancer, diabetes, neurodegenerative diseases, autoimmune and other immune dysfunctions, and viral disease; biotechnology products; veterinary medicines for use as growth or yield enhancers; and innovative products for advanced therapies, such as gene-therapy, somatic cell-therapy, or tissue-engineered medicines. Orphan drugs must also use the centralized procedure. The timeline for this procedure is 210 days.

Each European Union state member has its own national authorization system for determining drug approvability. Consequently, drug products that fall outside the scope of the centralized procedure can be granted marketing authorizations under national procedures. The national procedure allows the applicant to obtain a marketing authorization in only one member state. The timeline for this procedure is 210 days.

Under the decentralized procedure, an applicant can go directly to a national marketing authority and obtain marketing authorizations in that member state, and then seek to have other member states accept the marketing approval of the first member state. This allows the applicant to apply for authorization in more than one EU country simultaneously for products that have not been authorized in any EU country. The timeline for this procedure is 210 days.

The mutual recognition procedure allows an applicant to obtain marketing authorizations in a member state other than the member state where the drug product had previously received a marketing authorization. The timeline for the other member states to recognize the authorization of the first member state is 90 days.[95] The whole evaluation process may consume a time period of 390 days.

TABLE 38.8 Differences Between FDA and EMA Procedures Used for Drug Registration

FDA	EMA
One agency	Multiple agencies
	European medicines agency (EMA)
	Committee for Medicinal Products for Human Use (CHMP) of the EMA—scientific input
	National health agencies
One registration process	Multiple registration processes
	Centralized (all EU state members)
	National (one state member)
	Decentralized (at least two state members)
	Mutual recognition (at least two state members)
FDA reviewers within same agency	EMA just an administrative framework, national agencies provide the scientific reviewers
Specialized review of data	Overall review of benefit/risk of entire data
The changes in the approved drug can be done by filing PAS, CBE-30, CBE, annual reports	The changes in the approved drug can be done by filing Type IA, IB, and II variations

As summarized in Table 38.8, there have been differences between the United States and European Union in terms of procedures used for drug registration. Furthermore, the labeling format used to describe a drug product indication is different in the United States and European Union; whereas the FDA uses the United States Package Insert (USPI), the EMA uses the Summary of Product Characteristics. A study comparing labeling approved by the FDA and EMA showed notable differences in adverse events, black box warnings, order of information, number of approved indications, dosing regimen, and number of clinical trials described.[96]

Perhaps the most visible difference between FDA and EMA policies is regarding biosimilars; whereas interchangeability is defined by statute in the BPCI Act in the United States, in the European Union, its assessment resides with the member states. A relatively recent EMA draft guidance on biosimilar monoclonal antibodies may be used as a source of guidance for the structure and conduct of biosimilarity clinical studies for biosimilar applications in the European Union.[97]

Alignment between the FDA and EMA is increasing as a result of coordination and cooperation between these two agencies. One example of such alignment efforts between FDA and EMA is the pilot for QbD applications, which established a platform between the two agencies and allowed US and EU reviewers/assessors to exchange their views on ICH Q8. These ongoing efforts will eventually provide a streamlined procedure for drug approval and will lead to decreased costs for drug development and marketing.[50]

Acknowledgments

The authors would like to thank Dr Thomas Schultz (Johnson & Johnson), Dr Mohan Ganapathy (Merck), and Dr Emanuela Lacana (FDA) for critical review of this manuscript and their many helpful comments and suggestions.

References

1. *About FDA: FDA's Origin and Functions*. <http://www.fda.gov/AboutFDA/WhatWeDo/History/Origin/default.htm>; 2009 as of July 17, 2015.
2. <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/TherapeuticBiologicApplications/Biosimilars/default.htm>; as of July 17, 2015.
3. *Regulatory Information: Food and Drug Administration Safety and Innovation Act (FDASIA)*. <http://www.fda.gov/RegulatoryInformation/Legislation/FederalFoodDrugandCosmeticActFDCAct/SignificantAmendmentstotheFDCAct/FDASIA/ucm20027187.htm>; 2015 as of July 17, 2015.
4. *About FDA: FDA Organization*. <http://www.fda.gov/AboutFDA/CentersOffices/default.htm>; 2015 as of July 17, 2015.
5. *About FDA: ORA Overview*. <http://www.fda.gov/AboutFDA/CentersOffices/OfficeofGlobalRegulatoryOperationsandPolicy/ORA/ucm409371.htm>; 2014 as of July 17, 2015.
6. *About FDA: Initiatives at CDER*. <http://www.fda.gov/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/ucm410539.htm>; 2014 as of July 17, 2015.
7. *FDA Pharmaceutical Quality Oversight – One Quality Voice*. <http://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/UCM442666.pdf>; 2015 as of July 17, 2015.
8. *Pharmaceutical Quality for the 21st Century*. <http://www.fda.gov/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/ucm128080.htm>.
9. *FDA Pharmaceutical Quality Oversight – One Quality Voice*. <http://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/UCM442666.pdf>; 2015 as of July 17, 2015.
10. Sood R. GPhA CMC Workshop; June 9, 2015.
11. Iser R. *Update on Question-based Review (QbR)*, presented at Regulatory Education for Industry (REdI) Generic Drugs Forum, Silver Spring. <http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/SmallBusinessAssistance/UCM445619.pdf>; April 22–23, 2015 as of July 17, 2015.
12. CDER MAPP 4000.4 Clinical Pharmacology and Biopharmaceutics Review Template. <http://www.fda.gov/downloads/AboutFDA/ReportsManualsForms/StaffPoliciesandProcedures/ucm073007.pdf>; 2004 as of July 17, 2015.
13. *Chemistry Review of Question-based (QbR) Submission*. <http://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/ManualofPoliciesProcedures/UCM423752.pdf>.

14. Flynn MM. *Introduction to CMC review: question-based review (QbR) and risk-based review (RbR)* (2015). New Employee Training. FDA.
15. *FDA clears first 3D-printed drug*. <www.medscape.com/viewarticle/848971> as of October 14, 2015.
16. *The FDA's Drug Review Process: Ensuring Drugs are Safe and Effective*. <http://www.fda.gov/Drugs/ResourcesForYou/Consumers/ucm143534.htm>; 2014 as of July 17, 2015.
17. Troetel WM. The investigational new drug application and the investigator's brochure. In: Guarino A, editor. *New drug approval process*. New York, NY: Dekker; 2004.
18. *FDA website Drugs Frequently Asked Questions on Prescription Drug User Fees (PDUFA)*. <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/SmallBusinessAssistance/ucm069943.htm#P47_2024>.
19. *FDA website Orange Book Nomenclature*. <http://www.fda.gov/ohrms/dockets/ac/05/briefing/2005-4137B1_07_Nomenclature.pdf>.
20. *Frequently Asked Questions About Therapeutic Biological Products*. <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/TherapeuticBiologicApplications/ucm113522.htm>; 2015 as of July 17, 2015.
21. *New Drug Application (NDA)*. <http://www.fda.gov/drugs/developmentapprovalprocess/howdrugsaredevelopedandapproved/approvalapplications/newdrugapplicationnda/default.htm>; 2014 as of July 17, 2015.
22. *Investigational New Drug (IND Application)*. <http://www.fda.gov/drugs/developmentapprovalprocess/howdrugsaredevelopedandapproved/approvalapplications/investigationalnewdrugindapplication/default.htm>; 2015 as of July 17, 2015.
23. *FDA MAPP 6030.1 IND Process and Review Procedures*. <http://www.fda.gov/downloads/AboutFDA/ReportsManualsForms/StaffPoliciesandProcedures/ucm082022.pdf>.
24. 21 CFR 312.23 IND Content and Format.
25. FDA. *Guidance for industry content and format of investigational new drug applications (INDs) for phase 1 studies of drugs, including well characterized, therapeutic, biotechnology-derived products*; November 1995.
26. 21 CFR 314.50 Content and Format of an Application, Form FDA-356(h) Application to Market a New Drug, Biologic, or an Antibiotic Drug for Human Use. <http://www.fda.gov/downloads/AboutFDA/ReportsManualsForms/Forms/UCM082348.pdf>; 2014 as of July 17, 2015.
27. 21 CFR 201.56 (General) and 201.57 (Specific) Final Rule: requirements on the Content and Format of Labeling for Human Prescription Drug and Biological Products. <http://www.fda.gov/OHRMS/DOCKETS/98fr/06-545.pdf>; January 24, 2006, as of July 17, 2015.
28. FDA. *Guidance for industry guideline for the format and content of the summary for new drug and antibiotic applications*; February 1987.
29. 21CFR 314.50 (c) Content and Format of an Application.
30. FDA. *Guidance for industry drug product chemistry, manufacturing, and controls information*; January 2003.
31. Miksinski SP. Regulatory assessment of QbD elements in NDA submissions – experience and considerations, DIA 2013 49th annual meeting. <http://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/UCM406884.pdf> as of June 25, 2015.
32. ICH Q8(R2). <http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q8_R1/Step4/Q8_R2_Guideline.pdf>.
33. ICH M3(R2). *Guidance on nonclinical safety studies for the conduct of human clinical trials and marketing authorization for pharmaceuticals*. <http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002720.pdf>; 2009 as of June 4, 2015.
34. FDA 21 CFR 314.50. <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?fr=314.50>.
35. FDA. *Guidance for industry guideline for the format and content of the microbiology section of an application*. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm075101.pdf>; prior to February 1997 as of June 25, 2015.
36. *Drugs: clinical/medical*. <http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm064981.htm>; 2015 as of July 17, 2015.
37. FDA. *Guidance for industry guideline for the format and content of the clinical and statistical sections in an application*. <http://www.fda.gov/downloads/Drugs/Guidances/UCM071665.pdf>; July 1988 as of June 4, 2015.
38. FDA. *Guidance for industry submitting debarment certification statement*. <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm080584.pdf>; Draft, 1998 as of June 4, 2015.
39. *PDUFA user fee cover sheet*. <http://www.fda.gov/ForIndustry/UserFees/PrescriptionDrugUserFee/ucm119184.htm>; 2014 as of July 17, 2015.
40. FDA. *Guidance for clinical investigators, industry, and FDA staff financial disclosure by clinical investigators*. <http://www.fda.gov/downloads/RegulatoryInformation/Guidances/UCM341008.pdf>; 2013 as of July 17, 2015.
41. FDA. *Guidance for industry M4Q: the CTD-Quality*. <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm073280.pdf>; 2001 as of July 17, 2015.
42. FDA. *Guidance on providing regulatory submission in electronic format – certain human pharmaceutical product applications and related submission using the eCTD specification*. <http://www.fda.gov/ucm/groups/fdagov-public/@fdagov-drugs-gen/documents/document/ucm333969.pdf>; 2015 as of July 17, 2015.
43. <http://www.fda.gov/downloads/Drugs/.../Guidances/ucm153222.pdf>.
44. FDA MAPP 6010.5. *Policy and procedures office of new drugs NDAs and BLAs: filing review issues*. <http://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/ManualofPoliciesProcedures/UCM081990.pdf> as of June 25, 2015.
45. *For industry: what information is provided to sponsors during the human drug product application review process*? <http://www.fda.gov/ForIndustry/FDABasicsforIndustry/ucm238030.htm>; 2015 as of July 17, 2015.
46. FDA. *CDER 21st century review process desk reference guide, NDA/BLA review process*. <http://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/ManualofPoliciesProcedures/UCM218757.pdf>; 2015 as of June 26, 2015.
47. 21CFR. 314.110 Complete response letter to the Applicant.
48. *Advisory committees: drugs*. <http://www.fda.gov/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/default.htm>; 2014 as of July 17, 2015.
49. FDA. Guidance for industry expedited programs for serious conditions – drugs and biologics. <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm358301.pdf>; 2014 as of July 17, 2015.
50. Yu LX, et al. Advancing product quality: a summary of the inaugural FDA/PQRI conference. *AAPS J* 2015;**17**(4). Available from: http://dx.doi.org/10.1208/s12248-051-9754-4.

51. *For industry: designating an orphan product: drugs and biological products*. <http://www.fda.gov/ForIndustry/DevelopingProductsforRareDiseasesConditions/HowtoapplyforOrphanProductDesignation/default.htm>; 2013 as of July 17, 2015.
52. *Qualifying for pediatric exclusivity under section 505(A) of the federal food, drug, and cosmetic act*. <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DevelopmentResources/ucm077915.htm>; 2013 as of July 17, 2015.
53. *Abbreviated New Drug Application (ANDA): Generics*. <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/AbbreviatedNewDrugApplicationANDAGenerics/default.htm>; 2015 as of July 17, 2015.
54. *Question-based Review for CMC Evaluations of ANDAs*. <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/AbbreviatedNewDrugApplicationANDAGenerics/ucm120971.htm>; 2014 as of July 17, 2015.
55. 21CFR 320.22. *Criteria for waiver of evidence of in vivo bioavailability or bioequivalence*.
56. Yu LS, Li BV, editors. *FDA Bioequivalence standards*. New York, NY: Springer; 2014.
57. *The biopharmaceutics classification system (BCS) guidance*. <http://www.fda.gov/aboutfda/centersoffices/officeofmedicalproductsandtobacco/cder/ucm128219.htm>.
58. FDA. *Guidance for industry ANDA submissions – content and format of abbreviated new drug applications*. <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm400630.pdf>; 2014 as of July 17, 2015.
59. FDA. *Guidance for industry controlled correspondence related to generic drug development*. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM411478.pdf>; 2014 as of July 17, 2015.
60. FDA. *ANDA filing checklist*. <http://www.fda.gov/downloads/Drugs/UCM320405.pdf>; 2014 as of July 17, 2015.
61. *Generic drug user fee act program performance goals and procedures*. <http://www.fda.gov/downloads/ForIndustry/UserFees/GenericDrugUserFees/UCM282505.pdf> as of July 17, 2015.
62. *Drugs: generics*. <http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm064995.htm>; 2015 as of July 17, 2015.
63. Ya N. *Generic drugs- application and regulatory review*. <http://www.asq509.org/ht/a/GetDocumentAction/i/86200> as of July 17, 2015.
64. FDA. *Draft guidance, bioequivalence studies with pharmacokinetic endpoint for drugs submitted under an ANDA*; 2013.
65. Uhl K. *Creating the path to success with GDUFA, GPhA annual meeting*. <http://www.fda.gov/downloads/ForIndustry/UserFees/GenericDrugUserFees/UCM434884.pdf>; February 10, 2015.
66. FDA. *Guideline for drug master files (DMF)*. <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/DrugMasterFilesDMFs/ucm073164.htm>; 1989 as of June 4, 2015.
67. *Drug master files (DMFs)*. <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/DrugMasterFilesDMFs/default.htm>; 2015 as of July 17, 2015.
68. *Drugs drug master files: guidelines*. <http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm122886.htm>; 2014 as of July 17, 2015.
69. *FDA/CDER SBIA Chronicles: DMF 101*. <http://www.fda.gov/downloads/drugs/developmentapprovalprocess/smallbusinessassistance/ucm438678.pdf>; 2014 as of July 17, 2015.
70. *For industry: generic drug user fee amendments of 2012*. <http://www.fda.gov/ForIndustry/UserFees/GenericDrugUserFees/default.htm>; 2015 as of July 17, 2015.
71. Holland-Moritz. *Moving Biotech Products from CBER to CDER: A Work in Progress. Pharmaceutical Regulatory Guidance Book*. July 2006p. 70–3.
72. *Guidance for industry: scientific considerations in demonstrating biosimilarity to a reference product*. <http://www.fda.gov/downloads/DrugsGuidanceComplianceRegulatoryInformation/Guidances/UCM291128.pdf>; 2015 as of July 17, 2015.
73. *Drugs: biosimilars*. <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/TherapeuticBiologicApplications/Biosimilars/default.htm>; 2015 as of July 17, 2015.
74. Chow SS. *Introduction. Biosimilars design and analysis of follow-on biologics*. Chapman and Hall, CRC Press Taylor & Francis Group (multiple locations world-wide); 2013ISBN 9781466579699
75. *Guidance for industry: quality considerations in demonstrating biosimilarity of a therapeutic protein product to a reference product*. <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm291134.pdf>; 2015 as of July 17, 2015.
76. *Guidance for industry: biosimilars: questions and answers regarding implementation of the biologics price competition and innovation act of 2009*. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM444661.pdf>; 2015 as of July 17, 2015.
77. *Guidance for industry: formal meetings between the FDA and biosimilar biological product sponsors or applicants*. <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm345649.pdf>; 2013 as of July 17, 2015.
78. *Guidance for industry: clinical pharmacology data to support a demonstration of biosimilarity to a reference product*. <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm397017.pdf>; 2014 as of July 17, 2015.
79. *Purple book: lists of licensed biological products with reference product exclusivity and biosimilarity or interchangeability evaluations*. <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/TherapeuticBiologicApplications/Biosimilars/ucm411418.htm>; 2015 as of July 17, 2015.
80. *Guidance for industry: changes to an approved NDA or ANDA*. <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm077097.pdf>; 2004 as of July 17, 2015.
81. *Guidance for industry: established conditions: reportable CMC changes for approved drug and biologic products*. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM448638.pdf>; 2015 as of July 17, 2015.
82. *Guidance for industry: changes to an approved NDA or ANDA*. <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm077097.pdf>; 2004 as of July 17, 2015.
83. FDA. *Guidance for industry changes to an approved application: biological products*. <http://www.fda.gov/downloads/biologicsbloodvaccines/guidancecomplianceregulatoryinformation/guidances/blood/ucm170166.pdf>; 1997 as of July 17, 2015.
84. *Guidance to industry: changes to an approved application for specified biotechnology and specified synthetic biological products*. <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm124805.pdf>; 1997 as of July 17, 2015.
85. Cartwright AC. The common technical document – a global format for registration files. In: Cartwright AC, Matthews BR, editors. *International pharmaceutical product registration*. New York, London: CRC, Informa Healthcare; 2009. ISBN 9781420081763.
86. Jimenez AG. ICH and EU regulatory framework and the role of the European Medicines Agency (EMA), ASEAN Training Workshop on ICH Q5C, 30–31 May 2011, Kuala Lumpur; 2011.

87. Lezotre PL. *International cooperation, convergence, and harmonization of pharmaceutical regulations – a global perspective*. London: Elsevier; 2013.
88. *Process of harmonization*. <http://www.ich.org/about/process-of-harmonisation.html> as of July 17, 2015.
89. ICH Guidelines. <http://www.ich.org/products/guidelines.html> as of July 17, 2015.
90. *Quality guidelines/work products*. <http://www.ich.org/products/guidelines/quality/article/quality-guidelines.html> as of July 17, 2015.
91. *ICH Vision*. <http://www.ich.org/about/vision.html>.
92. *Guidance for industry: submitting marketing applications according to the ICH-CTD format general considerations*. <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm073308.pdf>; 2001 as of July 17, 2015.
93. *European medicines agency what we do*. <http://www.ema.europa.eu/ema/index.jsp?curl=pages/about_us/general/general_content_000091.jsp&mid = WC0b01ac0580028a42>; 2015 as of July 17, 2015.
94. Howie L, et al. A comparison of FDA and EMA drug approval: implications for drug development and cost of care. *Oncol J Pract Policy* 2013;**27**(12):1198–200.
95. *The mutual recognition procedure. European commission*. <http://ec.europa.eu/health/authorisation-procedures-mutual-recognition_en.htm> as of July 17, 2015.
96. Bordoloi JK, et al. *Comparison of the US package insert and the EU summary of product characteristics*. <http://pharmafellows.rutgers.edu/resources/df72777093f4adc706d3047e47c292b9.pdf> as of July 17, 2015.
97. EMA. *Guideline on similar biological medicinal products containing monoclonal antibodies – non-clinical and clinical issues*. <http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/06/WC500128686.pdf>; 2012 as of July 18, 2015.

This page intentionally left blank

CHAPTER

39

Modern Pharmaceutical Regulations: Quality Assessment for Drug Substances*

H. Zhang, D. Johnson, D. Skanchy and L.X. Yu

U.S. Food and Drug Administration, Silver Spring, MD, United States

39.1 INTRODUCTION

Information on the chemistry, manufacturing, and controls (CMC) for the drug substance must be submitted to support the approval of original new drug applications (NDAs) and abbreviated new drug applications (ANDAs) (21 CFR 314.50[d][1] and 514.1[b]). This chapter elaborates the assessment of the CMC information for drug substances in a structure corresponding with common technical document (CTD) format,[1] preferably as an electronic CTD (eCTD).[2] Under section 745A(a) of the Federal Food, Drug, and Cosmetic Act (FD&C Act), NDA, ANDA, and master files must be submitted in eCTD format beginning May 5, 2017, and certain investigational new drug submissions must be submitted in eCTD format beginning May 15, 2018.[3]

The CTD format arose from the efforts in the International Conference on Harmonization (ICH) to harmonize the varying documentation requirements that Japan, Europe, and the United States had for the submission of technical data to support the registration of pharmaceuticals for human use. The ICH Guidance for Industry "M4: Organization of the CTD" describes a harmonized format for the CTD that will be acceptable in all three regions. The CTD is organized into five modules. Module 1 is region-specific and contains administrative information, while Modules 2-5 are intended to be common for all regions. The guidance, "the CTD for the Registration of Pharmaceuticals For Human Use: Quality-M4Q(R1)", describes the format for the quality information submitted in Module 3 of an application and the CMC information for drug substance is specified in 3.2.S. 1-7. In Module 2, the overall summary of the technical data has a corresponding subsection 2.3.S for quality overall summary (QOS) for drug substance as well. The QOS is a summary of the information provided in Module 3, and therefore would not contain information, data, or justification that was not already included in that module (ICH M4Q).

This chapter discusses the quality assessment of drug substance with focus on the question-based review (QbR) format. The QbR for quality assessment started with the FDA's generic program in early 2005 and was implemented in 2007.[4] It is a concrete and practical implementation of the underlying concepts and principles outlined by the FDA's current good manufacturing practices (cGMPs) for the 21st Century and transforms the CMC review into a modern, science- and risk-based pharmaceutical quality assessment. These principles also apply to the assessment of drug substance CMC information that is not submitted using QbR. The QbR serves two purposes for the CMC assessment. First, it provides an assessment guide in the evaluation of whether a product is of high quality and determination of the level of risk associated with the manufacture and design of the product. Second, it provides transparency to sponsors about the logic behind a CMC review.

*__Disclaimer__: The chapter reflects the views of the authors and should not be construed to represent the views or policies of the Food and Drug Administration (FDA). No official endorsement by the FDA is intended or should be inferred.

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00039-X

39.2 ORIGIN OF THE QbR

The rapid growth of the generic industry has presented the agency with several challenges; namely, a significantly increased workload due to the high number of submissions, the vastly increased number of drug substances and drug product manufacturers with widely varied manufacturing experience, poor-quality CMC information being submitted, and new requirements to complete reviews in a more timely manner. In order to address these challenges, various changes have been implemented. One of these changes was the development of the QbR for quality assessment of ANDAs.[5] The QbR model is a general framework for a science- and risk-based assessment of product quality which is based on a series of focused questions divided into three broad categories: (1) drug substance quality standards, (2) drug product quality standards, and (3) process understanding and proposed drug manufacturing scale-up plans. Under the QbR review model for generic drug applications, ANDA applicants provide their responses to the questions in the QOS. The QbR model is not intended to replace the detailed supportive information in CTD Module 3.

QbR has benefited both industry and the FDA.[6] QbR provides a clear indication about what information needs to be provided to the agency. For the FDA, the sponsor-prepared QOS provides a quick overview of the entire CMC package and reduces the review time spent on documentation. It also improves the quality and consistency of the CMC information, which may lead to a reduction in the number of questions issued by the agency and a decrease in the review cycles. QbR provides an opportunity for applicants to demonstrate process understanding by addressing quality-by-design (QbD)[7]-related questions regarding the product's design, performance, manufacturing controls, and risk mitigation. QbD is a scientific, risk-based, holistic, and proactive approach to pharmaceutical development that focuses on deliberate design effort from product conception through commercialization and emphasizes a full understanding of how product attributes and the process relate to product performance. Knowledge gained through the pharmaceutical development studies and manufacturing experience provides scientific understanding of the process, which can be used to support the establishment of a design space. A design space is the multidimensional combination and interaction of input variables (eg, material attributes) and process parameters that have been demonstrated to provide assurance of quality. Upon approval, adjustments of processing parameters within an approved design space are not considered as a change and do not normally require a regulatory filing. This allows the sponsor more flexibility to execute manufacturing processes for which they have demonstrated process understanding.[8] The QbR program was fully implemented in 2007 and is currently being used by the generic industry for almost all ANDA submissions.

39.3 EVOLUTION OF THE DRUG SUBSTANCE REVIEW PROCESS

An application consists of information about both the drug substance, also known as API, and the drug product. The drug substance information may be supplied in one of two ways: (1) submitted directly in the application in 2.3.S of the QOS and 3.2.S of Module 3, or (2) submitted in a Type II Drug Master File (DMF). This type of DMF is a confidential, detailed document submitted by API manufacturers to the FDA containing the CMC information for a drug substance. Although the DMF is never actually approved, it is referenced by an application and the content is reviewed for scientific adequacy to support the action taken on the referencing application.[9]

Historically, the CMC sections and any associated DMFs were reviewed by a single discipline of chemistry. In Fall 2010, a pilot program was started in the FDA's generic program with a dedicated review team whose sole focus was the review of drug substance information found in DMFs. The 6 month pilot was very successful and led to the creation of the DMF Review Staff in early 2011.[10] The rationale for the change was to concentrate subject matter experts in the fields directly related to drug substance manufacture, such as organic synthesis, process engineering, and analytical chemistry. Another goal was to increase the consistency and efficiency of the review process by allowing the ownership of the DMF (ie, the assignment of DMF would stay in the same assessment team) until the submission is found adequate. The group has become the focal point for all drug substance quality issues and facilitates consistent and transparent communication with the DMF holders. Formation of the DMF Review Staff has allowed the agency to develop initiatives designed to improve the submission and review quality of the DMFs and has facilitated the implementation of the Generic Drug User Fee Act of 2012 (GDUFA).[11] The group has grown considerably with GDUFA hiring and eventually transformed into the Division of Lifecycle API as part of the Office of Pharmaceutical Quality (OPQ) reorganization, which is discussed next. A separation of drug substance and drug product review has also been implemented for the review of NDAs as part of the reorganization.

In parallel to the GDUFA implementation, OPQ, a new CDER super office, was strategically organized

in order to streamline the quality regulatory review processes. With a motto of "One Quality Voice," OPQ embodies the concept of putting the patient first by balancing risk and availability, as well as closer integration of review, inspection, surveillance, policy, and research activities for the purpose of strengthening pharmaceutical quality on a global scale.[12] A team-based integrated quality assessment (IQA) is reflective of the structure and function of OPQ, which brings together the field and review functions and their respective knowledge bases. This new approach makes the regulatory process more effective in the assessment of product quality and fosters the advancement of a continuous quality improvement-oriented culture.

The team-based IQA approach maximizes each team member's expertise and provides aligned patient-focused and risk-based quality recommendations for the drug substance, drug product, manufacturing process, and facilities.[13] IQA teams consist of an application technical lead (ATL), a regulatory business process manager (RBPM), discipline reviewers, and additional technical advisors as needed. Review disciplines may include drug substance, drug product, process (generally focused on drug product), facility, microbiology, biopharmaceutics, and field investigators. Technical advisors may come from OPQ laboratories, policy, surveillance, and other offices as needed. The ATL is responsible for overseeing the scientific content of the assessment, while the RBPM is responsible for driving the review process and adhering to the established timeline. The drug substance review function is consolidated into the Office of New Drug Product. The drug substance reviewers will assess the CMC information of drug substances whether it is submitted directly in the applications or in referenced DMFs.

The success of QbR and the separation of the drug substance review into the DMF Review Staff prompted another initiative, which was the development of a QbR template for API DMF review. With the reorganization to form OPQ, the QbR initiative became a joint effort among offices and divisions that deal with the drug substance quality assessment and eventually led to the recent publication of MAPP 5015.10, "Chemistry Review of QbR Submissions".[14] During development of MAPP 5015.10, the QbR was revised to improve the questions based on the lessons learned over the past several years and to encourage more incorporation of QbD elements in the CMC information submitted to the FDA.[15] While the revised QbR for drug product quality assessment will be discussed in the next chapter of this book, the following section elaborates the drug substance quality assessment.

39.4 QUALITY ASSESSMENT FOR DRUG SUBSTANCES

As mentioned previously, a QbR template for drug substance quality assessment aims to standardize the documentation practice and clarify the FDA's expectation in drug substance quality assessment. It follows the same format as the initial QbR developed for ANDAs and contains the same underlying QbD principles.

In May 2012, ICH Q11[16] was finalized; this guidance encourages manufacturers to use an enhanced approach that focuses on the use of "risk management and scientific knowledge to identify and understand process parameters and unit operations that impact critical quality attributes (CQAs) and develop appropriate control strategies applicable over the lifecycle of the drug substance which may include the establishment of design space(s) (ICH Q11)." This greater understanding of the drug substance and its manufacturing process can create the basis for more flexible regulatory approaches in a manner similar to that described in ICH Q8 for drug product. The questions developed for the drug substance QbR allow the manufacturers to demonstrate process understanding and justify control strategies other than end-product testing.

QbR is designed with the expectation that the API information is organized in CTD format and preferably submitted in eCTD format, which will become a requirement by May 2017.[3] The QOS (Module 2) follows the scope and outline of Module 3 and information provided in the QOS should be consistent with the information provided in Module 3. MAPP 5015.10 provides guiding principles for drug substance quality assessment, and can also serve as a guide for preparing the submission of drug substance information. Attachment 3 in the MAPP 5015.10 elaborates how the information will be evaluated under QbR format, but also applies to submissions that do not follow the QbR format. The companion document has many practical suggestions and considerations along with numerous example tables. In the following subsections of this chapter, additional details regarding QbR for drug substance quality assessment are discussed that extend beyond the companion document. This supplemental information may provide further insights, which can be considered in the submission and during assessment.

S.1 General Information

The general information section covers nomenclature, structure, and general properties. The nomenclature may include one or more of the following: International Nonproprietary Name, United States Approved Name, Compendial Name (eg, United States Pharmacopeia,

USP), Chemical Abstracts Service registry number, or Chemical name (using International Union of Pure and Applied Chemistry, IUPAC, nomenclature principles).

A chemical structure with molecular formula and molecular weight is commonly provided, including the stereo configuration if applicable. The presence of hydrates/solvates/salts is shown in the chemical structure, and the molecular weights for both active moiety and the salt or solvate/hydrate form are reported. When the API is a salt and the impurities are controlled at a weight/weight percent limit with respect to the API salt, attention should be paid to the calculation of the maximum daily dose. If the strength of the drug product was expressed as the active moiety in the drug product labeling, a conversion of the strength in active moiety to the strength in the salt form may be needed.

General properties may include a list of physicochemical properties, as described in the companion document, particularly those which are likely to impact the product performance and/or process design. For example, polymorphism/pseudopolymorphism and aqueous solubility characteristics may be CQAs to most solid oral dosage forms; hygroscopicity and light sensitivity may affect material handling procedure and storage configuration. The pK_a is an intrinsic property of the API molecule, but could be a factor in developing an efficient acid/base extraction procedure during API manufacture. Some of these properties may be further discussed in and referenced in section S.3: Characterization.

S.2 Manufacture

ICH M4Q does not set up subsections in Module 2, but the MAPP 5015.10 separates the questions into a subsection-format similar to the structure in Module 3.2. S; the next discussion also follows this structure for consistency.

S.2.1 Manufacturer

When addressing what facilities are involved in the manufacturing and testing activities, the following types of API-related facilities would be included in the submissions as applicable:

- API manufacturers
- API intermediate manufacturers
- API micronizers
- Facilities that sterilize the API
- Facilities that perform routine commercial testing of the API
- For animal/plant-derived APIs, the facilities that perform crude extraction prior to purification of API

It is very important to submit a complete set of facility information in the application so that all the facilities involved can be evaluated for compliance purpose and appropriate action (eg, approval or complete response) can be taken in a timely manner. A good manufacturing practices (GMP) compliance statement would be provided for each of the applicable facilities listed previously. It is less of an issue when the applicant is also the API manufacturer and well aware of what API-related facilities are involved. When manufacture of the API is contracted out or the API information is referenced to a separate submission (eg, a DMF) and the manufacturing process involves multiple sites and/or testing laboratories, it becomes critical that the applicant and the contractor or the holder communicate the facility information so that there is no missing facility in the application. A common issue that may lead to missing facility information involves starting material redesignation. For example, when an initially proposed starting material designation was not accepted, an API manufacturer may agree to redesignate a starting material in an earlier stage of the process, but these additional steps of the process may be performed at a separate site. In addition to providing the necessary CMC information for the process starting from the newly designated starting material, the DMF should also be updated with the new facility information in order for the new site to be evaluated before an action can be taken on the application.

S.2.2 Description of the Manufacturing Process and Controls

The information on the API manufacturing process typically includes route of synthesis, flow diagram, and detailed process description. A complete synthetic scheme would be the most effective way to illustrate the process, including molecular formulae and weights, chemical structures of starting materials and intermediates, the use of reagents, catalysts and solvents, and their equivalencies and relative ratio in each step. Nonisolated intermediates would normally be noted with brackets. The process is further described in the narrative process description, which should include batch size, input quantities, and yields. Some other operations, such as blending, micronization, reprocessing, or reworking should also be described as applicable, though they may not be performed routinely or may be subject to the customer requirement from time to time. When blending of small batches to increase batch size is needed, each batch for blending should have been tested against the established criteria; out-of-specification batches cannot be blended for the purpose of meeting the specification. If optional micronization or any specific operation is performed to meet a certain particle size requirement, such operation should be described and the development studies which demonstrate that such operation does not negatively impact the API quality and stability can be presented in Section 2.6 with cross reference.

When process analytical technology (PAT)[17] is proposed for routine production that allows for real-time process monitoring and control, the technology and the traditional method being replaced should be described and a discussion of how this technology will be implemented and the impact on overall control should be provided. The FDA's PAT initiative started in the early 2000s and establishes a framework to facilitate innovation with the goal of enhanced understanding and control of the manufacturing process. Process analysis has advanced significantly during the past several decades due to an increasing appreciation for the value of collecting process data. Available tools have evolved from those that predominantly take univariate process measurements, such as pH, temperature, and pressure, to those that measure chemical and physical attributes. At-line measurement refers to cases where the sample is removed, isolated from, and analyzed in close proximity to the process stream; on-line measurement is where the sample is diverted from the manufacturing process, and may be returned to the process stream; measurement where the sample is not removed from the process stream at all is in-line analysis. A description of these PAT tools and the traditional method being replaced should be provided. A discussion about how the technology will be implemented and the impact of this technology on the overall control strategy (ie, it is medium impact if it would be used for monitoring; it is high impact if it would be used for drug substance release) should be provided.

S.2.3 Control of Materials

Designation of regulatory starting material(s) is discussed in several FDA guidance documents and ICH guidelines[18] and the agency's official adoption of ICH Q11 in Nov. 2012 set a milestone in the regulatory assessment of the starting material designation.[19] Designation of starting material(s) in an NDA would typically be discussed at a pre-NDA stage. For an ANDA or its referenced DMF, the designation of starting material is assessed during the scientific review of the submission. A recently published "RTR guidance"[20] makes the appropriate designation of starting material(s) as one of the filing criteria. The guidance states that "FDA will refuse-to-receive (RTR) an ANDA if the API review (during the filing stage), whether in an ANDA or in a referenced DMF, reveals that the starting material for the API is not justified according to the principles in the ICH Q11 guidance." The inappropriate designation of regulatory starting materials represents an omission and inadequacy in the submission that impairs meaningful regulatory assessment of the API manufacturing process and quality control strategy, and would therefore render the application incomplete on its face and result in the failure of completeness assessment (CA)[21] in the case of an API DMF and possible RTR of the ANDA.

The assessment on the justification of starting material designation focuses on the overall quality control strategy, which manages and mitigates the risk associated with changes in the starting material source and/or the manufacturing process. The synthetic route and brief process description of the starting material can greatly facilitate the evaluation of the starting material control. Discussion on the relevant impurities arising from the manufacturing process of each proposed starting material and the ability of analytical procedures to detect impurities in the starting material will help justify the proposed specification of the starting material. Additional discussion on the fate and purge of those impurities and their derivatives in subsequent processing steps can address the Q11 general concept that impurities in the starting material will not have an impact on the impurity profile of the drug substance; this will further tie the starting material quality attributes to the overall control strategy.

When multiple starting material suppliers are being utilized, it is important that the starting material specification and overall control strategy are designed to address all of the various types of impurities that may arise in each synthetic route. In addition, when multiple manufacturing sites are involved during the manufacture of an API, the starting material specification(s) should ensure the same level of quality regardless of which site or starting material supplier is involved. Data demonstrating that starting material sourced from each current vendor has equivalent quality should also be provided as part of the justification of the specification for the starting material.

The first principle for designating an appropriate starting material in ICH Q11 recognizes the relationship between risk and the number of steps from the end of the manufacturing process; the second principle states that enough of the drug substance manufacturing process should be described in the application for regulatory authorities to understand how impurities are formed in the process, how changes in the process could affect the formation, fate, and purge of impurities, and why the proposed control strategy is suitable for the drug substance manufacturing process; and the third principle points out that manufacturing steps that impact the impurity profile of the drug substance should normally be included in the manufacturing process. It is also worth noting that a step, as defined in the glossary of ICH Q11, refers to a chemical transformation and typically involves C–X or C–C bond formation or breaking. Often in the justification of a regulatory starting material, the principle of "significant structural fragment"

is misinterpreted. "Significant structural fragment" in this context is intended to distinguish starting materials from reagents, solvents, or other raw materials. Justification of a late intermediate as a starting material by claiming it is a significant structural fragment is not considered a valid argument because it could apply to any intermediate in the manufacturing process. Based on these principles, selecting a regulatory starting material very late in the synthetic route may not be a viable submission strategy.[22]

If redesignating a regulatory starting material to an earlier stage of the process is deemed necessary, amending the application or its referenced DMF to address this issue may require a substantial amount of information. This process could be very involved and lengthy, especially if the originally proposed starting material is sourced externally. It is important to note that such a redesignation request cannot be simply addressed by only amending the submission with additional CMC information or referencing a separate DMF for the CMC information without actually declaring a new regulatory starting material in an earlier stage of the process because designation of regulatory starting material also has an implication in GMP compliance. Compliance with the cGMP is required from the introduction of the regulatory starting material to the manufacturing process.

Again, the general principles laid out in ICH Q11 should be considered in totality in selecting starting materials, rather than applying each general principle in isolation.

S.2.4 Control of Critical Steps and Intermediates

The intermediates addressed here are the ones that are isolated during the process so that intermediate specifications can be established and documented to ensure that quality attributes are met. The isolation of intermediates, with their corresponding specifications, provides great opportunities to purge potential impurities that may be carried from the starting materials or earlier stages. This is especially important when regulatory starting materials are in later stages of the process and/or are outsourced from multiple suppliers. The specifications should be designed to control the critical material attributes (CMAs) that may have an impact on the CQAs of a drug substance because intermediates are considered to be critical materials in API production and are thus a particularly important part of the overall control strategy.

When an API manufacturer chooses to outsource part of the GMP process, it is important to remember that GMP begins with introduction of the regulatory starting material(s) as mention in the previous subsection. A harmonized control strategy governing the entire GMP process would ensure the consistency of quality regardless of where the intermediate is produced. Batch data demonstrating that intermediates produced at all sites have comparable quality should also be provided, preferably in tabulated format. Hold-time studies need to be considered when intermediates have been outsourced and shipping between sites is required. These studies would demonstrate that the intermediate is stable in its packaging for the amount of time elapsed between manufacture and use. Hold-time studies are also appropriate for manufacturing processes, which require multiple batches of one intermediate to be combined in the next step of the process.

Potential impurities at a given intermediate stage include residual starting materials and previous chemicals, impurities carried over from the starting materials, reagents, solvents, catalysts, and reaction by-products. Investigation of impurities that are commonly observed is a key step in determining what impurities need to be controlled. It is essential for manufacturers to demonstrate process understanding by examining the origins and fates of impurities to support the proposed controls at various stages. ICH Q11 describes and recognizes that CQAs of a drug substance can be confirmed through upstream controls. Both in-process control (IPC) and impurity tests in the intermediate specification can be part of the upstream controls. When the limit for a particular impurity in the intermediate is set at a level higher than what would be acceptable in the API, inclusion of a study demonstrating how the amount of that particular impurity decreases throughout the subsequent manufacturing process from the proposed limit to an acceptable level would provide a stronger justification than submission of testing data in the final API alone and may warrant no further control in the later stages. Sometimes an impurity may not need to be specified in the final API specification if a significant margin of process purging capacity is demonstrated by spiking the impurity at a level much higher than the level it is commonly observed/generated at the intermediate stage and confirming its absence or amounts well below the acceptable level in the subsequent stage or final API. It is also important to consider whether an impurity may be transformed in later steps. Analysis of the API should also include potential by-products that may be formed due to transformation of the initial impurity into other by-products. Spike/purge studies should be conducted in a manner representative of the production process so that the outcome can predict the manufacturing capability.

S.2.5 Process Validation and/or Evaluation

Validation of the manufacturing processes is a requirement of the cGMP regulations for finished pharmaceuticals (21 CFR 211.100 and 211.110), and is considered an enforceable element of cGMP practice

for APIs under the broader statutory cGMP provisions of section 501(a)(2)(B) of the Federal Food, Drug, and Cosmetic Act. ICH Q7A[23] states that prospective validation should normally be performed for all API processes and should be completed before the commercial distribution of the final drug product manufactured from that API. ICH Q11 also indicates that the drug substance manufacturing process should be validated before commercial distribution of resulting drug product. However, none of the guidance specifically defines the submission requirement with regard to API process-validation information. The FDA Guidance to Industry, "Process Validation: General Principles and Practices",[24] does not specify what information should be included as part of a regulatory submission and refers interested persons to contact the appropriate center in determining the type of information to include in a submission. ICH M4Q mentions that process validation and/or evaluation studies for aseptic processing and sterilization should be included. ICH Q11 mentions as an observation that for nonsterile chemical entity drug substance processes, results of process validation studies are not normally included in the dossier. The "initial completeness assessments for Type II API DMFs under GDUFA"[21] is the only guidance that specifically requires "[a] summary of Process Validation and/or Evaluation information" to be submitted for a validated process (item 34 in its "Appendix 1: GDUFA Initial Completeness Assessment Checklist for Type II API DMFs").[21]

The scope of this validation summary is significantly smaller than the typical process validation report and does not involve the facility, utility, environmental control, equipment installation/operation/performance qualification (IQ/OQ/PQ), cleaning validation, etc., and analytical validation is covered in a separate section of the submission (S.4.3). Per the companion document to MAPP 50115.10, if the validation has been performed, information regarding the validation batches with lot traceability should be provided. The information should include the analytical data from the starting material(s), intermediates, in-process testing, and final drug substance, which can be provided in a tabulated format; reference to sections 2.3, 2.4 and 4.4 for the same analytical data is acceptable if the validation data is provided in those sections respectively. If the validation has not been completed, a validation commitment including the timeline and proposed scale can be provided.

S.2.6 Manufacturing Process Development

The process development summary should summarize the development effort, but should not be filled with raw study reports. It presents the results of the process development effort using summary charts and tables and some data can be cross-referenced to other sections as needed. A QbD approach or QbD elements are highly recommended to justify the choice of critical process parameters (CPPs), testing limits and ranges, and as part of risk mitigation in scale-up. There should be a discussion of how the proposed controls mitigate the identified risks in the process. The process development summary presents the opportunity to reveal the overall control strategy that the manufacturer has in place to ensure process robustness and consistent product quality.

One of the key elements of the control strategy is the impurity control, including controls of related substances, impurities with genotoxic structural alerts, reaction by-products, residual solvents, metal catalysts, and any impurities that may be carried from the starting material(s), etc. The manufacturers could demonstrate their process understanding by examining the origins and fates of impurities, discussing the probabilities and severities of their impacts to process robustness and the product quality, and conducting tests and/or purging studies to support the controls implemented at various stages.

In the case of chiral drug substances, discussion regarding how the chirality of each center is established through the synthetic process is very important. The possibility of interconversion of the chiral drug substance to the enantiomer or other diastereomers during manufacturing or storage, and the controls in place to prevent such interconversions also need to be addressed. A similar evaluation should be conducted for drug substances for which geometrical isomers are possible.

Regarding polymorphism, it should be demonstrated that the desired polymorphic form is consistently produced and retained during storage. The impact of specific process parameter(s) (eg, concentration, temperature program, choice of solvents and/or antisolvent, stirring/seeding as needed, etc.) should be discussed, and polymorphism analysis results should be provided. When there is a potential for other forms to be present, a discussion about how these forms are detected and controlled should be included. When the polymorphic form of the API is known in the literature, a reference to the patent or other literature can be provided.

The particle size of an API may impact product performance depending on the drug product dosage form, route of administration, and the product manufacturing process. If micronization or any specific particle size reduction operation is performed to meet the particle size requirement, the manufacturer should demonstrate that such operation does not negatively impact the quality and stability of the processed material. For example, complete release testing of batches could be performed after the customized

operations and a micronized sample should be put in the stability program.

S.3 Characterization

Characterization of the API should be performed irrespective of whether the drug has a compendial monograph or it is a new chemical entity as its purpose is to prove the API identity from a given process. A list of spectroscopic studies as applicable (eg, infrared [IR], ultraviolet [UV], nuclear magnetic resonance [NMR], mass spectroscopy [MS], x-ray diffraction [XRD], elemental analysis) and the interpretation of those studies summarizing the evidence of structure would be needed to adequately prove the structure. When the drug substance contains one or more chiral centers, it should be indicated whether it is a racemate or a specific stereoisomer. In the case of a specific stereoisomer, evidence of its configuration should be provided when an enantiomer or diastereomers are possible. For example, unequivocal characterization could be demonstrated by using 2D NMR spectroscopies and/or x-ray crystallography, or comparison with authentic samples or extracted reference listed drug in the case of generic or 505b(2) applications, using spectroscopic and/or chromatographic procedures that can discriminate between enantiomers or diastereomers. Alternatively, the stereochemistry can be established and controlled in the starting material or an intermediate when appropriate (eg, chirality feature inherited from the starting material or introduced in an intermediate stage). However, it should be demonstrated that subsequent process steps do not involve the chiral center(s) or won't cause its racemization. When the stereochemistry is established at an earlier stage, similar characterization should be performed at that point.

Potential genotoxic impurities (PGIs) are receiving more and more attention. Any work or studies performed to evaluate PGI's (eg, functional group identification, *in silico* analysis, Ames test, literature references, etc.) should be provided. However, it should be noted that a nonclinical consult review may be required to evaluate *in silico* and toxicological data and full copies of any study reports (ie, Ames studies, *in silico* analysis, etc.) are needed in those cases. A brief discussion about the control strategy for these impurities (ie, type of analytical procedures and sensitivity limits, in-process tests, tests in intermediate or final drug substance specifications, etc.) should be provided. A reference to process development summary may be appropriate, and the exact location of the information in S.2.6 should be provided.

The evaluation and justification of the proposed limits for an impurity with a structural alert is based on the relevant guidance, particularly the recently published ICH M7[25] and a stepwise approach can be taken to address the concern. A computational toxicology assessment can be performed using (quantitative) structure-activity relationships [(Q)SAR] methodologies that predict the outcome of a bacterial mutagenicity assay. Two (Q)SAR prediction methodologies that complement each other should be applied. The absence of structural alerts from two complementary (Q)SAR methodologies (expert rule-based and statistical) is sufficient to conclude that the impurity is of no mutagenic concern and no further genotoxicity assessments would be necessary. If warranted, the outcome of any computer system-based analysis can be reviewed using expert knowledge to provide additional supportive evidence on relevance of any positive, negative, conflicting. or inconclusive prediction and provide a rationale to support the final conclusion.

If the (Q)SAR prediction of potential mutagenicity is positive or the (Q)SAR study is not conducted, either adequate control measures at the threshold of toxicological concern (TTC)-based limits could be applied or a bacterial mutagenicity assay with the impurity alone can be conducted.[26,27] An appropriately conducted negative bacterial mutagenicity assay would overrule any structure-based concern, and no further genotoxicity assessments would be recommended. These impurities would be considered nonmutagenic; therefore, the impurity control can be based upon ICH Q3A(R2) recommendations.[28,29]

A positive bacterial mutagenicity result would warrant further hazard assessment and/or control measures (TTC-based limit). If compliance with the TTC-based limits is impossible or impractical, an in vivo gene mutation assay should be conducted to understand the relevance of the bacterial mutagenicity. Results from an appropriate in vivo assay may support setting compound specific impurity limits. In cases where control efforts cannot reduce the level of the mutagenic impurity to below the acceptable limit and levels are already as low as reasonably practicable (ALARP), a higher limit may be justified based on a risk/benefit analysis. Conversely, the application of ALARP is not necessary if the level of the mutagenic impurity is below acceptable limits.

S.4 Control of Drug Substance

This section addresses the control of the drug substance with respect to specification, analytical procedures and their validation, batch analyses, justification of specification, and overall control strategy. A specification as defined in ICH Q6A[30] is a list of tests, references to analytical procedures, and appropriate acceptance criteria for those tests, which are expressed as numerical limits, ranges, or other criteria as applicable. The drug substance must conform to this set of

criteria in order to be considered acceptable for its intended use. The specification is proposed and justified by the manufacturer and must be acceptable to regulatory authorities.

The drug substance manufacturer is encouraged to include relevant controls in the specification, but the specification is only one part of the total control strategy. As mentioned previously, ICH Q11 discusses an enhanced approach which focuses on the use of "risk management and scientific knowledge to identify and understand process parameters and unit operations that impact CQAs." Table 39.1 is from ICH Q11; it shows an example of an enhanced control strategy, which may be used by a holder to justify the use of upstream control to ensure consistent product quality.

Notice that upstream control can be a test for an impurity in the starting material or intermediate, or an in-process test. Upstream control may also be assured by process development data that demonstrates the process is able to adequately remove the impurity, in which case there is no routine testing involved. In either case, it is important that the manufacturer provide adequate information to justify their proposed control strategy. As discussed in the section S.2.4, if upstream control is established at higher than the acceptable limit or where control is achieved solely through process capability, spike/purge studies at greater than the proposed limit, and/or sufficient batch data from drug substance manufactured from starting material/intermediate containing the proposed levels of the impurity should be supplied to support the proposed limit. It is important that spike/purge studies be performed in a manner representative of the commercial process in order for the results to be predictive. A description of the analytical procedures used for these studies with information about limit on detection/limit on quantification (LOD/LOQ) should also be included.

A commonly observed issue in this section is related to the specification justification; particularly inadequate justification of proposed impurity limits.[31] For a new drug substance, the impurity justification typically goes by the threshold-based approach described in ICH Q3A. Additional justification may include the relevant development data, test data for the batches used in the toxicology and clinical studies, the data from the stability studies can be referred to as well as; reasonable variability of manufacturing process and analytical procedure should also be considered; other factors (eg, patient population, drug class effects, and clinical considerations) may be considered on a case-by-case basis.

For a drug substance subject to an ANDA, a guidance for industry has been published on how to appropriately justify impurity limits.[32] The qualification threshold in ICH Q3A should be the default

TABLE 39.1 Example of a Possible Control Strategy Summary

	Type of control			
DS CQA/ Limit	**In-process controls (including in-process testing and process parameters)**	**Controls on material attributes (raw materials/ starting materials/ intermediates)**	**Impact of manufacturing process design**	**Is CQA tested on drug substance/included in drug substance specification?**
Impurity X (NMT[a] 0.15%)	Design space of the reflux unit operation composed of a combination of %water in intermediate E and the reflux time in step 5 that delivers intermediate F with hydrolysis impurity ≤0.30% (3.2.S.2.2)			Yes/yes
Impurity Y (NMT 0.20%)	Process parameters step 4 (3.2.S.2.2) p(H2) ≥2 barg; T <50°C; In-process test step 4 (3.2.S.2.4) Impurity Y ≤0.50%			Yes/yes
S-enantiomer (NMT 0.50%)		Spec for starting material D (3.2.S.2.3), S-enantiomer ≤0.50%	Stereocenter is shown not to racemise (3.2.S.2.6)	No/no
Toluene	In-process test step 4 (3.2.S.2.4) ≤2000 ppm by GC		Process steps after step 4 are shown to purge toluene to levels significantly below (less than 10%) that indicated in ICH Q3C (3.2.S.2.6)	No/no[b]

[a]*NMT, not more than.*
[b]*This approach could be acceptable as part of a control strategy when justified by submission of relevant process data that demonstrates the adequacy of the process design and control. The manufacturing process should be periodically evaluated under the holder's quality system to verify removal of the solvent.*
Reproduced from: ICH Q11.

starting point for a specified impurity. An impurity is considered qualified when it meets one or more of the following conditions:

- The observed level and proposed acceptance criterion for the impurity do not exceed the level observed in the reference listed drug product
- The impurity is a significant metabolite of the drug substance, as confirmed by supporting data from studies or literature
- The observed level and the proposed acceptance criterion for the impurity are adequately justified by the scientific literature
- The observed level and proposed acceptance criterion for the impurity do not exceed the level that has been adequately evaluated in toxicity studies

Lack of proper justification of impurity limits may lead to RTR the application.[33]

Descriptions of the analytical procedures referenced in the specification are also reported in S.4 and the QbR companion document provides example tables that may be useful in presenting the information. There are generally two types of analytical procedures: in-house methods, which are developed by the applicant or holder, and compendial methods, which are contained in a monograph or a general chapter. In both cases, a description of the method should be provided. The method validation/verification reports for these methods may be included in this section or submitted in regional section. Full validation reports should be provided for in-house methods. If the original method validation was done at one site, such as a research and development (R&D) laboratory, and is then transferred to another site for routine release testing, a method transfer study should be conducted to verify the receiving site is capable of performing the analytical testing method for its intended use. When a compendial method is available and used, method verification should be performed. Verification under conditions of use is generally demonstrated by meeting system suitability criteria established for the method, as well as other method parameters such as: demonstration of specificity, method precision, and LOQ if applicable.[34] However, there are times when a supplemental validation of the compendial method should be performed. For example, if a compendial method is used to analyze unique in-house related substances, then validation of all parameters with regard to these impurities should be performed (eg, specificity, system/method precision, accuracy, linearity/range, robustness, etc.).

As for the batch analysis data submitted in this section of QoS, it is commonly tabulated against the specification. When the API is manufactured at more than one site, more data may be needed. As mentioned under S.2.1, adequate site information should be provided along with a GMP statement for alternative manufacturing sites. It is important to indicate when this site has previously received a satisfactory cGMP inspection and specify the type of manufacturing process covered by the inspection. The inspection status is one of the factors used to determine the post-approval filing category if the site change/addition is part of post-approval activity. Utilizing the same process in the alternative site would help to minimize the potential variation in drug substance quality, while it is acknowledged that changes to accommodate the facility and equipment difference may be necessary provided they result in no impact to the impurity profile, physiochemical properties, and control strategy. A tabulated comparison of the manufacturing process at each site with discussion/justification about any variation in equipment or process parameters is very useful in justifying changes and assists in the evaluation of their potential impact to drug substance quality. A detailed comparability study in impurity profile and physiochemical properties should be conducted and summarized to demonstrate the equivalence of the API manufactured at both sites. In addition, a summary of the plans for validation at the alternative site as elaborated under S.2.5, comparative data from the existing site should be provided and API from the alternative site should be added to the stability program.

A change in scale of the commercial batch size also requires additional information be submitted to the agency. A discussion about any variation in equipment, procedure, or control strategy caused by the scale change should be provided in the appropriate sections of the submission. A scale change, as referred to here, does not involve modification of the synthetic process. Comparative batch data between batches of API manufactured by the original process and batches manufactured by the scaled process should be provided to show that the quality is comparable.

S.5 Reference Standards

Per ICH Q6A Guideline, a reference standard, or reference material, is a substance prepared for use as the standard in an assay, identification, or purity test. Reference standards have both quantitative and qualitative uses. Reference standards for assay/purity tests for drug substances are for quantitative use, while reference standards for identification tests, system suitability tests, or chromatographic peak markers are for qualitative use. Other reference standards include performance standards and calibrators, such as dissolution calibrators, melting point standards, particle count set standards, chromatography calibration verification standards, etc. The amount of tests needed to qualify a reference standard depends on its intended use. For example, tests for the reference standards used for assay/purity tests typically include identification, assay, chromatographic purity, water, solvents, ash, and other tests as applicable. For primary standards the method of choice in

computing the assigned content value is mass balance analysis using independently determined components, such as moisture, solvent residues, inorganic residues (ash), chromatographic impurities, and ion contents.

If compendial reference standards are available and used, the compendial lot numbers would be considered sufficient. For noncompendial reference standards, source, batch/lot number, brief description of the preparation and qualification (characterization and quantification) of reference standards would be needed. The characterization data can be referenced to the data in S.3 if applicable. Quantitation is often done by means of mass balance. Certificate of analyses (COAs) of the in-house reference standards should include the storage conditions and expiry date.

The impurity reference standards, which are used in the release of a drug substance and/or validation of the analytical procedures, can be described in this section as well. Similar to the drug substance reference standards, the impurity reference standards would also need to be qualified to demonstrate their identities and purities/assays. Though often their purity values may not be on par with that of a drug substance reference standard, the impurity reference standards should be pure enough to not compromise the analysis. Normally, using a less-pure impurity reference standard may artificially inflate the reported impurity amount than its actual level. Again, if compendial impurity reference standards are available and used, the compendial lot numbers would be sufficient.

Manufacturers often utilize secondary reference standards (or working reference standards) for routine analysis in order to preserve the precious primary reference standards, which may have been highly purified and comprehensively characterized. Once identity and assay of the primary reference standard are fully established, the secondary reference standards can be qualified against the primary reference standard with a minimum of identity and assay tests. Qualification of a reference standard should not be confused with release of drug substance because the intended purpose of a reference standard is different from drug substance. The latter is used for manufacturing drug product, which would ultimately be dosed to patients; and the former is used in an analytical laboratory and consumed or disposed during or after the physical and chemical tests. Therefore, some tests required for releasing drug substance lots may not necessarily apply to the reference standard qualification; though a reference standard lot may have received a full release testing when it is sourced from a production lot.

S.6 Container Closure System

APIs are generally solids, and the container closure system for storage or shipment of a bulk solid APIs is typically a drum with double low density polyethylene (LDPE) liners that are usually heat-sealed or closed with a twist tie. A desiccant may be placed between the bags. The drum provides protection from light and mechanical strength to protect the liner during shipment and handling. The majority of the protection from air and moisture is provided by the liner. An API that is moisture-sensitive may need additional protection because LDPE is not a particularly good moisture barrier. An alternative to a LDPE bag is a heat-sealable laminate bag with a comparatively low rate of water vapor transmission.

The container closure system for the storage or shipment of a bulk liquid drug substance is typically a container with a rugged, tamper-resistant closure. Qualification of the container closure system may include characterization for solvent and gas permeation, light transmittance, closure integrity, ruggedness in shipment, protection against microbial contamination through the closure, and compatibility and safety of the packaging components as appropriate (eg, plastic, stainless steel, a glass-lined metal container, or an epoxy-lined metal). More information regarding general consideration on qualification and quality control of packaging components is discussed in the Container Closure Systems Guidance.[35]

In this section, a description of the container closure system for the bulk API as well as the composition of the specific components (container, closure, all liners, inner seal, and desiccant [if any]) can be provided. Documentation from each vendor showing compliance for each primary packaging material with 21 CFR (174-186) (food safety statement) is typically considered sufficient to establish the safety of the materials of construction.[36] Source(s) and the Certificate(s) of Analysis would demonstrate that the appropriate qualification and quality control of packaging components is in place.

A summary of how the container closure system ensures proper product quality during shipping and storage, based on known properties of the drug substance (eg, hygroscopic, light or air sensitive, etc.), should also be submitted. The summary could include a description of how the packaging system used during stability studies is representative of the commercial packaging system. In addition, the container label could capture a description of any special storage conditions (eg, protection from light, humidity, etc.). The storage condition is preferably a numerical temperature range based on the stability studies, but using USP terminology associated with a defined temperature limit or range such as "USP Controlled Room Temperature" is often acceptable. Undefined general terms, such as "room temperature," should be avoided.

The container label is often provided in this section, or may sometimes be cross-referenced to the executed

batch record (EBR) or Regional Section. Among the information on the label, it is recommended that a retest/expiry date is included. Although the COA typically has this information, it is often stored/archived separately from the actual container. Having the retest/expiry date on the container label allows the operator, auditor, and/or investigator to conveniently verify that the material is still in the valid period for use. ICH Q7A indicates that the retest date should be indicated on the label and/or COA for APIs with a retest date. The FDA Compliance Program Guidance Manual (Program 7356.002F, Chapter 56) further stresses that "Adequate expiration or retest dates should be on the label."[37] In addition, 21 CFR 201.122 also recommends a caution statement, such as "caution: for manufacturing, processing, or repacking" and prescription sign "Rx" be on the label when most of the dosage forms in which the API may be used are subject to prescription.[38]

S.7 Stability

The recent publication of the ANDA stability guidance[39] largely harmonized the stability requirements among the new and generic drug applications to ICH Q1 guidelines.[40] A question and answer guidance was published in May 2014 to further clarify the FDA's recommendations regarding stability testing data for ANDAs and referenced API DMFs.[41] A unique situation pertaining to Type II API DMF under GDUFA often causes some confusion to the applicants and/or holders. To pass the CA, DMFs should include the stability protocol, commitments, and data demonstrating that stability studies have started. The initial and one additional time point for the accelerated studies and long-term studies are sufficient. The CA is strictly triggered by the GDUFA DMF fee payment even before any ANDA submission makes reference to it. The DMF holders are encouraged to pay the DMF fee earlier to initiate the CA process in order to bring the DMFs to the status of "available for reference." If the DMF does not meet the stability data requirement specified in the ANDA Stability Guidance at the time of the CA, the DMF holder should amend the DMF with updated stability data to prepare for full scientific review.

The stability acceptance criteria is often harmonized with the release criteria and the stability protocol design and the choice of API storage condition aligns well with intended storage and use condition of the drug product. Most of the drug products in solid oral dosage forms are stored at controlled room temperature. When the long-term stability studies are conducted at $5 \pm 3^{\circ}C$ and the API is intended for storage in a refrigerator while the corresponding drug product is or will be stored at the common controlled room temperature, the choice of refrigerated storage for the API should be explained. The following questions may be used as a guide when addressing such situations:

- Have the stability studies at room temperature longer than 6 months (typical 6 months of data at $25 \pm 2^{\circ}C/60\%$, RH $\pm$ 5% RH (relative humidity), as accelerated condition when long-term condition is $5 \pm 3^{\circ}C$) been conducted?
- Have the stability studies at elevated temperatures (eg, $30 \pm 2^{\circ}C/65\%$ RH $\pm$ 5% or $40 \pm 2^{\circ}C/75\%$ RH $\pm$ 5% RH) been conducted?
- Has any stability failure or trend been observed in those studies?
- How can the stability of a drug substance, which requires refrigerated conditions, be ensured in a drug product that is stored at controlled room temperature?

When some stability failure and/or clear trend of increasing degradants is observed, efforts should be made to identify and specify these degradants so that they can be continuously monitored throughout the product's shelf life. Additional information regarding these degradants as applicable may be needed to support the control strategy for them, such as:

- The chromatograms of impurity analysis at the failed stability time points
- The impurity profile tabulated with their RT/RRT and % w/w
- The trend analysis of the degradants to project their growth rate at room temperature
- Impurity characterization effort (or semi-characterization by LC/MS)
- Inclusion of the degradants that were observed with Out of Specification (OOS) results at room temperature in both release and stability specification (eg, by RRT and Mass ID if not fully identified yet) so that the drug product manufacturer(s) will know what to follow-up
- Qualification of the degradants(s) as needed per the applicable guidance to demonstrate that its exposure will not have additional risk to the patients
- In the case of the drug substance manufacturer being different from the drug product manufacturer, commitment to provide information about any observed stability failure, including results of the investigation and impurity characterization information, to the drug product manufacture(s)

39.5 CONCLUSION

QbR facilitates a high-quality submission with the appropriate level of detail. It standardizes submission

format, clarifies the FDA's expectations, provides an opportunity for applicants and holders to address critical questions about the product's design, failure risk, and manufacturing controls from both a performance and patient usability perspective, encourages justification for choices made throughout product development and manufacture, and increases transparency in the applicant's thought processes. It may also be used within the applicant's organization as an internal communication tool in preparation for the submission. The QbR also facilitates the team-based IQA in OPQ. It is consistent with the QbD paradigm, congruent with risk management approaches and also acts as a guide for assessing the level of risk associated with the design and manufacture of the product. This helps ensure a consistent and comprehensive quality evaluation and eventually leads to more focused and efficient review.

APPENDIX QbR QUESTIONS—DRUG SUBSTANCE[14]

2.3 Introduction to the Quality Overall Summary

2.3.S.1 General Information

1. What are the nomenclature, molecular structure, molecular formula, CAS number, molecular weight, and pharmacological class of the drug?
2. What are the physical, chemical, biological and, if applicable, mechanical properties including physical description, pK_a, chirality, polymorphism, aqueous solubility as a function of pH, hygroscopicity, melting point(s), and partition coefficient?

2.3.S.2 Manufacture

3. Who manufactures the drug substance? List each participant and facility involved in drug substance manufacturing/testing activities and clearly state their function. List the date of the last FDA inspection of each facility involved and the result of the inspection. Has the manufacturer addressed all concerns raised at the FDA inspection?
4. What is the flow diagram of the manufacturing process that shows all incoming materials, reagents, reaction conditions, and in-process controls and, if appropriate, any reprocessing/reworking/alternative processes?
5. If applicable, what on-line/at-line/in-line monitoring technologies are proposed for routine commercial production that allows for real-time process monitoring and control? Provide a summary of how each technology was developed.
6. What is (are) the starting material(s) for the manufacturing process and how would changes in starting material quality and/or synthesis/source be controlled to minimize adverse effects on the drug substance quality?
7. What are the starting material specifications and how are they justified?
8. What are the specifications for reagents, solvents, catalysts, etc.? What are the critical attributes for these materials that impact the quality of the final drug substance?
9. What are the critical process parameters (CPPs) and how are they linked to drug substance quality?
10. What are the in-process controls (IPCs)/tests associated analytical procedures and acceptance criteria for each control?
11. What are the specification(s) for the intermediate(s)?
12. What process validation and/or evaluation information is provided, if any?
13. What development and scale-up information supports the commercial process and control strategy?

2.3.S.3 Characterization

14. How is the drug substance structure characterized?
15. What are the potential impurities (eg, related substances, degradants, inorganic impurities, residual solvents) in the drug substance? Which of these impurities are potentially genotoxic?

2.3.S.4 Control of Drug Substance

16. What is the drug substance specification and what is the justification? Does the specification include all of the drug substance's critical quality attributes (CQAs)?
17. For each test in the specification, provide a summary of the analytical procedure(s) and, if applicable, a summary of the validation or verification report(s).
18. How do the batch analysis results compare with the proposed specification? Provide a summary of the batch analysis results.
19. What is the proposed control strategy for the drug substance manufactured at commercial scale? What are the residual risks upon implementation of the control strategy at commercial scale?

2.3.S.5 Reference Standards

20. How are the drug substance reference standards obtained, certified, and/or qualified?

2.3.S.6 Container Closure System

21. What container closure system(s) is proposed for commercial packaging of the drug substance and how is it suitable to ensure the quality of the drug substance during shipping and storage?

2.3.S.7 Stability

22. What are the stability acceptance criteria? If applicable, what is the justification for acceptance criteria that differ from the drug substance release specification?

23. What is the proposed retest period for the drug substance? What drug substance stability data support the proposed retest period and storage conditions in the commercial container closure system? How does statistical evaluation of the stability data, if any, and any observed trends support proposed the retest period?

24. What are the post-approval stability protocols and other stability commitments for the drug substance?

References

1. *ICH M4. Organisation of the common technical document for the registration of pharmaceuticals for human use.* <http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/CTD/M4_R3_Organisation/M4_R3__organisation.pdf>; [accessed 20.08.15].
2. *FDA Electronic Common Technical Document (eCTD) webpage.* <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/ElectronicSubmissions/ucm153574.htm>; [accessed 20.08.15].
3. *FDA Guidance for Industry. Providing regulatory submissions in electronic format—certain human pharmaceutical product applications and related submissions using the eCTD specifications.* <http://www.fda.gov/ucm/groups/fdagov-public/@fdagov-drugs-gen/documents/document/ucm333969.pdf>; [accessed 20.08.15].
4. *FDA QbR White Pater. Question-based review (QbR) for generic drugs: an enhanced pharmaceutical quality assessment system.* <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/AbbreviatedNewDrugApplicationANDAGenerics/ucm120973.htm>; [accessed 20.08.15].
5. Yu LX, Raw A, Lionberger R, et al. US FDA question-based review for generic drugs: a new pharmaceutical quality assessment system. *J Generic Med* 2007;**4**:239–48.
6. Iser R. Question based reviews. PDA/FDA Joint Technical Conference September 8–10, 2014. Washington, DC.
7. Yu LX, Amidon G, Khan MA, Hoag SW, Polli III J, Raju GK, et al. Understanding pharmaceutical quality by design. *AAPS J* 2014;**16**(4). <http://link.springer.com/article/10.1208%2Fs12248-014-9598-3#>. [accessed 20.08.15]
8. *ICH Harmonised Tripartite Guideline. Pharmaceutical development Q8(R2).* <http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q8_R1/Step4/Q8_R2_Guideline.pdf>; [accessed 20.08.15].
9. *Art Shaw. Introduction to the drug master file (DMF) webinar*. November 14, 2011. <http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/SmallBusinessAssistance/UCM279666.pdf>; [accessed 20.08.15].
10. Yu L. The future of DMF review in OGD. GPhA/FDA API workshop. October 1, 2012. Bethesda, MD.
11. *Food and Drug Administration Safety and Innovation Act (FDASIA).* Pub. L. No. 112–144, 126 Stat. 993. 2012. <http://www.gpo.gov/fdsys/pkg/BILLS-112s3187enr/pdf/BILLS-112s3187enr.pdf>; [accessed 20.08.15].
12. *FDA office pharmaceutical quality white paper.* <http://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/UCM442666.pdf> [accessed 20.08.15].
13. Yu L. GDUFA: evolving quality assessment. GPhA/FDA Fall Technical Conference October 28, 2014. <http://www.gphaonline.org/media//cms/Lawrence_Yu_2.pdf>; [accessed 20.08.15].
14. *Center for drug evaluation and research MAPP 5015.10. Chemistry review of question-based review (QbR) submissions.* <http://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/ManualofPoliciesProcedures/UCM423752.pdf>; [accessed 20.08.15].
15. Maguire J. Revision of QbR for QbD submissions. GPhA/FDA CMC Workshop. May 22–23, 2012. Bethesda, MD.
16. *ICH Harmonised Tripartite Guideline. Development and manufacture of drug substances (chemical entities and biotechnological/biological entities) Q11.* <http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q11/Q11_Step_4.pdf>; [accessed 20.08.15].
17. *FDA. Guidance for industry. PAT—a framework for innovative pharmaceutical development, manufacturing, and quality assurance.* <http://www.fda.gov/downloads/Drugs/Guidances/ucm070305.pdf>; [accessed 20.08.15].
18. Smith J. Drug substance starting materials and ICH Q11. ISPE Washington Conference. June 6–9, 2011. Washington, DC.
19. Skanchy D, Scott B. Current regulatory practices and expectations in drug master file review. Learning@AAPS Webinar. May 15, 2014. <http://www.aaps.org/eLearning/Webinars/2014/Current_Regulatory_Practices_and_Expectations_in_Drug_Master_File_Review/>; [accessed 20.08.15].
20. *FDA. Guidance for industry. ANDA submissions — refuse-to-receive standards.* <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm370352.pdf>; [accessed 20.08.15].
21. *FDA. Guidance for industry. Initial completeness assessments for Type II API DMFs under GDUFA.* <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM321884.pdf>; [accessed 20.08.15].
22. *European Medicines Agency. EMA/448443/2014. Reflection paper on the requirements for selection and justification of starting materials for the manufacture of chemical active substances.* <http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/10/WC500175228.pdf>; [accessed 20.08.15].
23. *ICH Harmonised Tripartite Guideline. Good manufacturing practice guide for active pharmaceutical ingredients Q7.* <http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q7/Step4/Q7_Guideline.pdf>; [accessed 20.08.15].
24. *FDA. Guidance for industry. Process validation: general principles and practices.* <http://www.fda.gov/downloads/Drugs/Guidances/UCM070336.pdf>; [accessed 20.08.15].
25. *ICH Harmonised Tripartite Guideline. M7. Assessment and control of DNA reactive (mutagenic) impurities in pharmaceuticals to limit potential carcinogenic risk.* <http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Multidisciplinary/M7/M7_Step_4.pdf>; [accessed 20.08.15].
26. Sutter A, Amberg A, Boyer S, Brigo A, Contrera JF, Custer LL, et al. Use of in silico systems and expert knowledge for structure-based assessment of potentially mutagenic impurities. *Regul Toxicol Pharmacol* 2013;**67**:39–52 <http://www.sciencedirect.com/science/article/pii/S0273230013000688>. [accessed 20.08.15].
27. *ICH Guideline S2(R1). Guidance on genotoxicity testing and data interpretation for pharmaceuticals intended for human use.* <http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S2_R1/Step4/S2R1_Step4.pdf>; [accessed 20.08.15].
28. *ICH Harmonised Tripartite Guideline. Impurities in new drug substances. Q3A(R2).* <http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q3A_R2/Step4/Q3A_R2__Guideline.pdf>; [accessed 20.08.15].
29. *OECD Guideline for Testing of Chemicals Test 471.* Bacterial reverse mutation test section 4. July, 1997 <http://www.oecd.org/chemicalsafety/risk-assessment/1948418.pdf>, [accessed 20.08.15].

30. *ICH Harmonised Tripartite Guideline. Specifications: test procedures and acceptance criteria for new drug substances and new drug products: chemical substances. Q6A.* <http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q6A/Step4/Q6Astep4.pdf>; [accessed 20.08.15].
31. Srinivasan A, Iser R. Common deficiencies in abbreviated new drug applications, Part 1: drug substance, pharmaceutical technology. *Pharm Technol* 2010;**34**(1):50–9. <http://www.pharmtech.com/fda-perspectives-common-deficiencies-abbreviated-new-drug-applications-part-1-drug-substance>. [accessed 20.08.15].
32. *FDA. Guidance for industry ANDAs: impurities in drug substances.* <http://www.fda.gov/downloads/Drugs/Guidances/UCM172002.pdf>; [accessed 20.08.15].
33. *FDA. Guidance for industry ANDA submissions—refuse to receive for lack of proper justification of impurity limits.* <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM414598.pdf>; [accessed 20.08.15].
34. *FDA. Analytical procedures and methods validation for drugs and biologics.* <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm386366.pdf>; [accessed 20.08.15].
35. *FDA. Guidance for industry*. Container closure systems for packaging human drugs and biologics. Chemistry, manufacturing, and controls documentation. <http://www.fda.gov/downloads/Drugs/Guidances/ucm070551.pdf>; [assessed 20.08.15].
36. *Food and Drug Administration*. Center for drug evaluation and research. MAPP 5015.5, Rev. 1. CMC Reviews of Type III DMFs for packaging materials. <http://www.fda.gov/downloads/AboutFDA/CentersOffices/CDER/ManualofPoliciesProcedures/UCM205259.pdf>; [accessed 20.08.15].
37. *FDA. A compliance program guidance manual (Program7356.002F, Chapter 56).* <http://www.fda.gov/downloads/iCECi/Compliancemanuals/ComplianceProgrammanual/ucm125420.pdf>; [accessed 20.08.15].
38. *21 C.F.R.Sect. 201.122. 2013.* <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfCFR/CFRSearch.cfm?fr=201.122>; [accessed 20.08.15].
39. *FDA. Guidance for industry. ANDAs: stability testing of drug substances and products.* <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm320590.pdf>; [accessed 20.08.15].
40. *ICH Harmonised Tripartite Guideline. Stability testing of new drug substances and products Q1A(R2).* <http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q1A_R2/Step4/Q1A_R2__Guideline.pdf>; [accessed 20.08.15].
41. *FDA. Guidance for industry. ANDAs: stability testing of drug substances and products.* Questions and Answers. <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm366082.pdf>; [accessed 20.08.15].

This page intentionally left blank

CHAPTER

40

Modern Pharmaceutical Regulations: Quality Assessment for Drug Products

J.A. Maguire, W. Jiang and L.X. Yu

U.S. Food and Drug Administration, Silver Spring, MD, United States

40.1 INTRODUCTION

Information on the chemistry, manufacturing, and controls (CMC) for the drug product must be submitted to the Food and Drug Administration to support the approval of original new drug applications (NDAs) and abbreviated new drug applications (ANDAs). This chapter discusses assessment of the CMC information for drug products with a focus on question-based review (QbR). The QbR is structured to facilitate the assessment of applications submitted in the International Conference on Harmonization (ICH) common technical document (CTD) format. Although QbR was initially introduced for drug products submitted in ANDAs, its principles apply equally well to drug products submitted in NDAs.

40.2 QbR HISTORY

In the early 2000s, pharmaceutical quality was redefined by the FDAs Pharmaceutical Current Good Manufacturing Practices (cGMPs) and quality-by-design (QbD) initiatives,[1] as well as ICH Q8(R2), Q9 and Q10,[2–4] and in response, applicants were encouraged to implement QbD approaches in their drug product development and manufacturing activities. Accordingly, the FDA Office of Generic Drugs (OGD) developed a new quality assessment system as a tool to enable assessment and promotion of QbD approaches in ANDAs (ie, QbR) (Fig. 40.1). However, OGD was not the first office within Center for Drug Evaluation and Research (CDER) to use a QbR assessment system. The initiative built upon the positive experiences of the Office of Clinical Pharmacology and Biopharmaceutics (OCPB), which implemented a standardized NDA review template in a question-based format to improve consistency, foster good interdisciplinary communication and teamwork, and lead the reader logically through the thought process used in resolving scientific, clinical, and regulatory issues.[5] QbR used by both OCPB and OGD are reflective of good review practices (GRPs), which are documented best practices within CDER. All GRPs share fundamental values that make them exceptional and require them to be disseminated and adopted by review staff. These fundamental values are quality, efficiency, clarity, transparency, and consistency.[6] Another impetus for the creation of QbR was feedback from other regulatory authorities (eg, Health Canada) that use the quality overall summary (QOS) as a foundation for the primary chemistry review document.[7] The development of QbR followed four underlying principles[8]:

1. quality is built into the product and process by design, development, and manufacture, and is confirmed by testing;
2. risk-based approaches should be used to maximize economy of time, effort, and resources;
3. best practices of the pre-QbR OGD review system and organization should be preserved; and
4. best available science and open communication with stakeholders should be utilized to facilitate implementation.

Under the QbR system, applications are submitted in CTD format and include a pharmaceutical development section, as outlined in ICH Q8R2 and a QOS in Module 2, that addresses the QbR questions. QbR results in a minimal change for applicants generating

Developing Solid Oral Dosage Forms
DOI: http://dx.doi.org/10.1016/B978-0-12-802447-8.00040-6

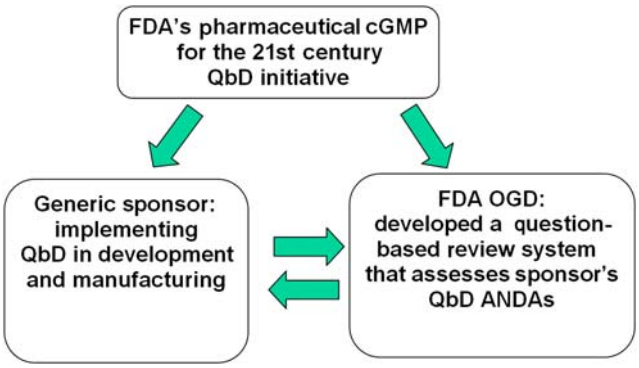


FIGURE 40.1 The Food and Drug Administration's cGMP initiative: quality by design, and question-based review.

multi-region submissions, so it does not impact the detailed submission package in Module 3. The Manual of Policies and Procedures (MaPP) 5015.10 on *Chemistry Review of Question-based Review (QbR) Submissions* clarifies how FDA reviewers should assess NDAs and ANDAs that follow a QbR format in conjunction with the FDA guidance for industry M4Q: The CTD—Quality.[9] Companion documents attached to MaPP 5015.10 detail the type of information expected within a submission in response to the QbR questions.[7]

For CMC reviews, a well-prepared QOS written by the sponsor provides the primary reviewer with a quick overview of the entire CMC package and reduces the review time spent on documentation. Pharmaceutical development information helps reviewers have a better understanding of how drug substance, formulation, and process variables affect the performance and stability of the drug product and supports appropriate performance-based specifications as well as the overall control strategy.

Due to its significant benefits, QbR was well-received by ANDA applicants and the FDA OGD chemistry reviewers, and has both parties' commitment. In Jul. 2007, more than 90% of ANDA submissions included a QOS that answered the QbR questions.[10] Today, 100% of ANDA applicants utilize QbR. Details about the impact and benefits of QbR are discussed in several commentary papers[8,11–18] and summarized in the MaPP 5015.10 and will not be repeated here.

40.3 CURRENT STATUS OF QbR

The recently published MaPP 5016.1 *Applying ICH Q8(R2), Q9, and Q10 Principles to CMC Review* establishes a policy that CMC reviewers will consider ICH Q8(R2), Q9, and Q10 recommendations when reviewing applications whether or not the applications include QbD approaches.[19] Reviewers will ensure that applications contain a quality target product profile (QTPP), critical quality attributes (CQAs) of the drug product, identification of those aspects of drug substances, excipients, container closure systems, and manufacturing processes that are critical to product quality and that support the safety and efficacy of the drug product, information that conveys an understanding of the development of the drug product and its manufacturing process, and a justified control strategy. Reviewers will take a scientific and risk-based approach during assessment of the application. The MaPP applies to both innovator and generic drugs.

CDER's quality initiative began more than a decade ago when the agency first introduced the 21st Century Initiative to modernize the FDA's regulation of the pharmaceutical quality of drugs.[1] The recent reorganization at CDER and the formation of the Office of Pharmaceutical Quality (OPQ) is a natural evolution of the review and inspection processes that allow the agency to better respond to an ever-changing world with continued globalization of the pharmaceutical industry, increasing product complexity, continued drug shortages and recalls, and legislative commitments outlined in the Prescription Drug User Fee Act (PDUFA) and the newly implemented Generic Drug User Fee Amendments of 2012 (GDUFA). OPQ further enhances CDER's quality initiative by creating a drug quality program as robust as the programs the agency already has in place for drug safety and efficacy. The formation of OPQ is a major step toward the goal of modernizing FDA regulation of drug quality and brings the CMC review of innovator and generic drug products into the same super office. The OPQ slogan is "One Quality Voice." The office integrates quality review, inspection, and surveillance over a drug product's lifecycle so that both internal and external stakeholders receive a single quality assessment capturing OPQ's overall recommendation on product approvability. OPQ emphasizes:

- A lifecycle approach to drug quality
- Proactive team-based and risk-based approaches
- Consistent and scientifically-sound quality standards for review and inspection
- Clear expectations and enforcement policies for industry
- Modernization and efficiency of drug manufacturing technologies

With widespread agreement that QbD and QbR have enhanced the quality of generic drugs, a recent initiative within OPQ focused on developing a QbR review template with overarching questions applicable to both innovator and generic drug products. The goals of the initiative were to standardize the review approach and expectations for both NDAs and ANDAs, better capture QbD expectations, incorporate

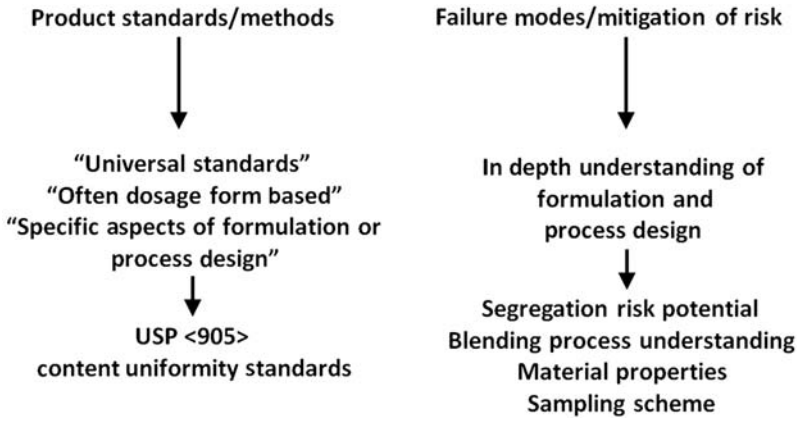


FIGURE 40.2 Conceptual dissection of a CMC review.

feedback and lessons learned from the previous QbR version implemented in 2007, support implementation of the integrated team-based review, which was introduced in chapter "Modern Pharmaceutical Regulations: Quality Assessment for Drug Substances," and facilitate communication among all quality stakeholders. The QbR questions presented in this chapter stemmed from this initiative and are covered in MaPP 5015.10 *Chemistry Review of Question-based Review (QbR) Submissions*.

The questions in the QbR recognize the two parts to a quality review that are conceptually different. Some questions reflect the need to comply with drug substance and drug product quality standards. These are universal standards that are often dosage-form based and address specific aspects of products, which are needed to deliver clinical performance. An example of a product standard is the expectation that all solid oral dosage forms include uniformity of dosage units in the drug product specification to comply with USP <905>. The USP acceptance criterion is a parametric two-sided tolerance interval test that covers 91% of the population at an 84% confidence level. This acceptance criterion ensures no more than 9% of the population will fall below 85% or exceed 115% at an 84% confidence level (ie, there is a 16% chance that a batch will be accepted with more than 9% out-of-specification). Results are generated on a sample population of up to 30 units, but batch sizes can range up to more than a million tablets or capsules. Therefore, it is debatable if passing results for such a minimal sample size is enough to demonstrate content uniformity for the entire batch. Thus, other questions go beyond compliance with general product standards and strive for a deeper, in depth understanding of the formulation and process design, probe the potential failure modes for the drug product, and focus on risk mitigation. These questions explore how risk is mitigated through reducing the probability and increasing the detectability of failures. For a low-dose drug product with a high risk of content uniformity failure, the focus is on minimizing segregation potential, optimizing mixing steps and demonstrating homogenous content with justified sampling schemes. The two parts of a quality review, product standards and risk mitigation, are illustrated in Fig. 40.2.

40.4 QbR QUESTIONS

QbR questions are designed to assess, concretely and practically, applicants' implementation of concepts and principles of QbD. At the heart of QbR are fundamental questions that reviewers ask themselves when assessing an application.

1. Will the product design ensure desired performance?
2. Will the applicant be able to scale-up to commercial scale and ensure comparable quality to the registration/exhibit batch(es)?
3. Will the applicant be able to manufacture the product with defined quality parameters over time?

The QbR questions for drug product are introduced next.

Introduction to the QOS

P.1 Description and Composition of the Drug Product

What is the description of the proposed commercial drug product? What are the components and composition of the final drug product as packaged and administered on both a per unit dose and %w/w basis? What is (are) the function(s) of each excipient?

The applicant should provide a descriptive summary of each strength of the drug product including weight, shape, dimensions, color, em/debossing, score configuration, etc. Any co-packaged components (eg, device components) should also be described. Often, a tabular

summary is an effective way of presenting this information. For generic drug products, submitting high-resolution photographs of all strengths of both the proposed generic and the reference listed drug (RLD) with a reference scale is also helpful. The composition table should be both qualitative and quantitative. All components of the drug product should be listed whether or not they appear in the drug product, including fluids used for granulation and gases used as a manufacturing aid. The subcomponents of any coating mixtures, colors, flavors, and inks should be provided. Where applicable, the composition table should clearly indicate the composition of significant drug product intermediates (eg, granules, tablet cores, beads). For each component, the quality standard and the function of that component in the formulation is needed. If a formulation contains a drug substance overage, this should be indicated in the composition table and justified in the submission.

Does any excipient exceed the FDA inactive ingredient database (IID) limit for this route of administration calculated based on maximum daily dose? If so, please justify.

This question asks for a comparison of each inactive ingredient level to the limit published in the electronic inactive ingredient database (eIID).[20] The comparison should be based on the maximum daily dose of the drug product and should be for the appropriate route of administration. Applicants should also discuss any patient safety considerations associated with the inactive ingredients when the drug product may be administered to special patient populations, such as pediatric or immunocompromised patients and/or when the drug product is meant to be administered chronically.

FDA will refuse-to-receive an ANDA if the submission proposes to use an inactive ingredient at a level that exceeds any of the eIID listings without adequate justification. In cases where the applicant wishes to use an inactive ingredient at a level per unit that is higher than what is proposed in the eIID, three options are available to facilitate receipt of the ANDA[21]:

1. Submit complete pharmacology/toxicology information
2. Cite a specific example of a CDER-approved drug product that contains the inactive ingredient at or above the proved level of use for the appropriate route of administration
3. Refer to an FDA-controlled correspondence response

If applicable, what are the differences between this formulation and the listed/RLD formulation?

This question applies specifically to 505(b)(2) applications for new drugs as well as 505(j) applications for generic drugs. Some ways in which a drug product submitted as a 505(b)(2) application may differ from the RLD include dosage form, strength, formulation, active ingredient, dosing regimen, or route of administration.[22] A generic drug product must be formulated to contain the same amount of active ingredient in the same dosage form and to meet the same compendial or other applicable standards (ie, strength, quality, purity, and identity) as the RLD.[23] A generic drug product should generally have the same score configuration as the RLD[21,24] and should be similar in size.[25] It should also have the same conditions of use; however, conditions of use for which approval cannot be granted because of exclusivity or an existing patent may be omitted. A generic drug product may differ in characteristics such as shape, color, and release mechanism.[23] With some exceptions, a generic drug product may also differ in excipients and drug substance solid state so long as safety and product performance will be equivalent to the RLD.

Any differences between the proposed generic drug product and the RLD should be justified in terms of therapeutic equivalence. Potential differences may include formulation, physical attributes, or design (eg, the generic drug product is designed as an extended-release matrix tablet, whereas the RLD is an extended-release osmotic tablet). For modified-release products, if a formulation with a different release mechanism from the RLD is used, the applicant needs to explain the differences, and provide comparative dissolution data between the RLD and the ANDA products in multiple dissolution media. In addition, it is possible for two products with very different plasma concentration profiles (such as different shapes or different T_{max}) to have equivalent AUC and C_{max}. The applicant needs to provide justification as to why the differences in plasma concentration profiles do not affect the therapeutic equivalence. Satisfactory answers to this question demonstrate scientific understanding of the underlying formulation design and biopharmaceutical principles.

P.2 Pharmaceutical Development

For 505(b)(1) applications, what is the rationale for selecting the proposed dosage form for the drug product? For 505(b)(2) and 505(j) applications, what are the characteristics of the listed/RLD product? What is the QTPP of the finished product based on the proposed indication and patient population? How is the QTPP justified?

This question solicits the scientific rationale behind the drug product design and encourages applicants to begin with the end in mind. New drug applicants should provide a brief description of the scientific and clinical rationale for the selected dosage form. Generic drug applicants should provide a complete characterization of the RLD product, including information regarding the clinical, pharmacokinetic, drug release, and physicochemical properties. Pharmacokinetic information such as AUC, T_{max}, and C_{max} is readily available

from the RLD labeling. The RLD labeling also provides critical information regarding alternative methods of administration. For example, some tablets may be crushed and suspended and some capsules may contain beads that can be sprinkled over applesauce. The applicant must bear these key design features in mind when developing a generic drug product that must be therapeutically equivalent to the RLD. Drug release properties should be investigated in several media/apparatus/dissolution conditions relevant to the specific dosage form. For solid tablets, physical characterization may include size, weight, shape, score configuration, and strength differentiation while chemical characterization may include dissolution profile, purity profile, stability, content uniformity, excipient grade and amount. The estimated composition of the RLD is usually based on the RLD labeling, patent literature, and/or reverse engineering. Another important consideration is the RLD packaging characterization, which focuses on what type of desiccant or O_2 scavengers are used.

Applicants should clearly define target product profile (TPP) and QTPP at the beginning of development. TPP is primarily expressed in clinical terms, such as clinical pharmacology, indications and usage, contraindications, warnings, precautions, adverse reactions, drug abuse and dependence, overdosage, etc. For a generic product, this information can be readily obtained from the labeling section of the RLD.[12] A generic drug product is therapeutically equivalent to the RLD in terms of dosage form, strength, route of administration, quality, performance characteristics, safety, and intended use. Therefore, a generic version and its RLD would be expected to have the same TPP.

The QTPP is a prospective summary of the quality characteristics of a drug product that ideally will be achieved to ensure the desired quality, taking into account safety and efficacy of the drug product.[4] The QTPP forms the basis of design for the development of a product that delivers the therapeutic benefit promised in the TPP.[12] Considerations for inclusion in the QTPP include the following[4,26]:

- Intended use in a clinical setting, route of administration, dosage form, and delivery system(s)
- Dosage strength(s)
- Container closure system
- Therapeutic moiety release or delivery and attributes affecting pharmacokinetic characteristics (eg, dissolution and aerodynamic performance) appropriate to the drug product dosage form being developed
- Drug product quality criteria (eg, identity, assay, content uniformity, purity, stability, drug release, sterility, and physical attributes) appropriate for the intended marketed product

Along with other available information from the scientific literature and possibly the pharmacopeia, the QTPP can be used to define product specifications, to some extent, even before the product is developed. Although it seems obvious that the target characteristics of a new product should be adequately defined before any development work commences, the value of predefining these target characteristics of the drug product has often been underestimated over the years, resulting in wasted time and valuable resources. A recent paper by Raw et al.[27] illustrates the significance of defining the correct QTPP before conducting any development. In addition, QbD examples exemplify the identification and use of QTPPs.[28–30] With a clear TPP and QTPP defined, formulation scientists can establish strategies and keep formulation efforts focused and efficient.

What are the quality attributes of the finished product? Which quality attributes are considered CQAs? For each CQA, what is the target and how is it justified?

For a solid oral dosage form, quality attributes may include identity, assay, content uniformity, degradation products, residual solvents, drug release or dissolution, moisture content, microbial limits, and physical attributes such as color, shape, size, odor, score configuration, and friability. These attributes can be critical or not critical. Identification of the drug product CQAs from the list of quality attributes in the QTPP is the next step in drug product development. A CQA is a physical, chemical, biological, or microbiological property or characteristic of an output material including finished drug product that should be within an appropriate limit, range, or distribution to ensure the desired product quality.[31] Criticality of an attribute is primarily based upon the severity of harm to the patient, should the product fall outside the acceptable range for that attribute. Probability of occurrence, detectability, or controllability does not impact criticality of the product quality attribute.

Drug product CQAs derived from the QTPP and/or prior knowledge are used to guide the product and process development. Risk assessment is used to identify the CQAs that have a high risk of failure based on characteristics of the drug product. The high-risk CQAs should be thoroughly evaluated during formulation and/or manufacturing development. A classic example of a CQA that typically has a high risk of failure is the content uniformity of a low-dose drug product. On the other hand, CQAs, such as identification, can be managed within the applicant's quality management system and are not required to be studied during pharmaceutical development because these CQAs generally have

a low risk of being affected by formulation or manufacturing variables.

The target for each CQA should be defined early in development based on the properties of the drug substance, dosing instructions, intended patient population, and characterization of the RLD product, as applicable. The target should be quantitative, where possible, and scientifically justified. During development, the target established in the QTPP may be modified or adjusted when the formulation and manufacturing process are selected and as product and process understanding is increased, as long as the rationale for the change is described.

What is the approach for meeting the CQAs related to clinical performance? If applicable, what in vitro bioperformance evaluations (ie, dissolution method, flux assay, etc.) were used during pharmaceutical development to ensure clinical performance?

A main goal of QbD is to achieve meaningful product quality specifications that are based on clinical performance. This question encourages early identification of the CQAs that are linked to efficacy and safety in the patient and development of in vitro tests to ensure clinical performance. The clinical effect may be direct or indirect depending on the pharmaceutical and toxicological determinants of the desired therapeutic performance. Such determinants of desired therapeutic performance can include therapeutic index, time to both onset and loss of clinical effect, duration of therapy (ie, acute or chronic), need for dose titration, washout or elimination period, pro-drug or metabolites, presence of toxic impurities or adventitious agents, inter and intrapatient variability, immunogenicity, and consequences of therapeutic failure. A nonexhaustive list of the CQAs that may be linked to these clinical determinants includes strength, purity, content uniformity, drug release or dissolution, and sterility. During the course of product development, the applicant should make every reasonable effort to use a dissolution test that is predictive of in vivo performance, especially for Biopharmaceutics Classification System (BCS) Class II (low solubility and high permeability) immediate-release drug products and for complex dosage forms such as extended release formulations. An acceptable response to this question would include a discussion of the relevance of the bioperformance test as a predictor of in vivo performance as well as a discussion of the method development and a description of any pilot bioavailability or bioequivalence studies performed to refine the method.

P.2.1 Components of the Product

The key objective of product design and understanding is to develop a robust product that can deliver the desired QTPP over the product shelf life. Product design is open-ended and may allow for many design pathways.[26] The QbR questions related to drug product design and understanding focus on the following key elements:

- Comprehensive characterization of the drug substance(s)
- Identification and selection of excipient type and grade, and knowledge of intrinsic excipient variability
- Interactions of drug substance and excipients
- Optimization of formulation and identification of critical material attributes (CMAs) of both drug substance and excipients

P.2.1.1 Drug Substance

What are the physical, chemical, biological and, if applicable, mechanical properties of the drug substance including physical description, pK_a, chirality, polymorphism, aqueous solubility (as a function of pH), hygroscopicity, melting point(s), and partition coefficient and, when available, BCS classification?

This question is similar to a drug substance QbR question presented in the previous chapter. With respect to the drug product, the physicochemical, biological, and mechanical properties of the drug substance that can influence the performance or stability of the drug product and its manufacturability, or the properties specifically designed into the drug substance (eg, solid state properties), should be identified and discussed. Details of these properties are presented in Table 40.1.[32,34] Solid understanding and characterization of drug substance properties are crucial to the successful development of a robust formulation and manufacturing process. Besides collecting information from literature and drug master file (DMF) sponsors, applicants are strongly encouraged to perform their own preformulation studies to obtain some key properties of drug substance, such as pH-dependant solubility and stability profile, solid state properties (eg, polymorphism), bulk and tapped density, etc. The impact of pharmaceutical solid polymorphism

TABLE 40.1 Typical Drug Substance Properties[32,33]

Physical and mechanical property	Color, odor, particle size distribution, shape, crystallinity, polymorphic form, hygroscopicity, bulk density, tapped density, true density, surface area, electrostatic charge, surface energy, elasticity, plasticity (ductility), viscoelasticity, brittleness
Chemical property	pK_a, aqueous solubility as a function of pH, chemical stability in solid-state and solution state, photolytic and oxidative stability
Biological property	Partition coefficient, membrane permeability, and/or oral bioavailability

and form transformation on product quality and performance has received much attention in recent years, and the findings have been presented in several review articles.[35–37] Decision trees that provide recommendations on monitoring and controlling polymorphs in drug substances and/or drug products can be referred to the recently published guidance.[38]

Based on the aqueous solubility, dose, and intestinal permeability of drug substance, the applicants may want to determine its BCS class.[34,39] With BCS class identified, applicants can understand and identify what may be the pertinent material attributes that can affect the drug product development and propose reasonable formulation strategies. For example, for BCS Class II drug substances, particle size reduction may be proposed as a formulation approach to improve delivery and manufacturability of a dissolution-limited drug substance. For a drug with 10 μg/mL solubility, in order for the drug product to achieve 80% dissolution in 30 minutes, the particle size has to be reduced to about 10 μm (assuming log-normal distribution, sink conditions, spheres).[40] Rohrs et al. published a paper to predict the effect of particle size distribution on content uniformity.[41] For a 1-mg dose, the particle size of drug substance, with very broad distribution (eg, geometric standard deviation is 3.5), is predicted to be about 15 μm to pass the USP <905> stage 1 content uniformity test for tablets, whereas for monodisperse particles (eg, geometric standard deviation is 1), particles around 150 μm are able to meet the content uniformity requirement. In these examples, particle size and distribution of drug substance are CMAs, which must be controlled to ensure product quality, and their specifications can be set based on understanding how they affect product performance.

In another example, oxidative degradation of the drug substance may require a stabilizer to be included in the formulation. It is key to understand the oxidation mechanisms of the drug substance, and then select appropriate stabilization strategies.[42,43] If the drug substance oxidation is heavy metal catalyzed, chelators such as ethylene diamine tetraacetic acid (EDTA) and citric acid should be used in the formulation. In cases where peroxides are believed to be the problem, stabilization strategies can be focused on careful selection of excipients with lower peroxide levels, lower concentrations of suspect excipients, and use of antioxidants.

What is the drug substance specification used to accept the incoming drug substance batches and how is it justified? For each test in the specification, provide a summary of the analytical procedure(s) and, if applicable, a summary of validation or verification report(s).

The drug substance specification should be inclusive of all quality attributes that may impact drug product manufacture and quality and may contain additional tests beyond those performed by the drug substance supplier. The specification is best presented in a tabular format with columns for the test, the acceptance criterion, the analytical method, and the full testing results for a representative lot(s) used in the registration batches. Although vendor qualification is the purview of the field, the applicant should distinguish tests that will be routinely performed by the applicant from those where the results will be taken from supplier's certificate of analysis (CoA). The applicant should also indicate what tests are performed during drug substance retesting. If complete IUPAC impurity names are not included in the table, then they may be presented as a footnote to the table. A separate table listing the impurity IUPAC name, structure and origin (process or degradant) is also helpful.

A summary of each noncompendial method should be provided. This can be in tabular or narrative form and should include the critical method parameters and the system suitability criteria. The applicant should also include a tabular summary of the method validation/verification performed per ICH Q2 and USP <1225> for all quantitative instrumental test methods (ie, assay, impurities, particle size, residual solvents, etc.) that includes the validation parameters evaluated, acceptance criteria, and results. Validation information is not necessary for USP methods or methods adopted from a DMF holder, but the suitability of the methods for their intended use should be verified (per USP <1226>) under actual operating conditions. For example, assay methods can be verified for specificity and precision and impurity methods can be verified for specificity, precision, linearity, and limit of detection (LOD)/limit of quantitation (LOQ).

When an in-house method or method from an alternative compendium (eg, European Pharmacopoeia (EP), British Pharmacopoeia (BP)) is chosen in lieu of the USP method, the results of a cross-over study to demonstrate that the method is equivalent or better than the USP method are needed. The comparison of analytical procedures is discussed in USP <1010>.

P.2.1.2 Excipients

What evidence supports excipient–drug substance compatibility and if applicable, excipient–excipient compatibility?

The inactive ingredients, or excipients, present in a formulation are themselves chemical moieties with functional groups that may be reactive. Sometimes there are unintended consequences or interactions including chemical degradation or complex formation between drug substance and excipient, loss of potency, degradation products with adverse events, or unacceptable changes in dosage form appearance or bioavailability.

QbD requires identification of mechanistic and formulation factors that affect drug product stability.

Acceptable finished drug product stability cannot be accepted as evidence of excipient–drug substance or excipient–excipient compatibility because end product testing only demonstrates that the quality was reached for the tested samples, but does not address how the product is designed to achieve the quality. Based on ICH Q8 recommendations, a drug substance/excipient compatibility study should be evaluated as a part of pharmaceutical development.[4]

Typically, a drug substance/excipient compatibility study is performed with binary mixtures or prototype formulations, which are prepared in the presence/absence of added water to mimic process steps (eg, wet granulation), stored at accelerated stability or stress testing conditions, and then analyzed by stability-indicating methodology, such as high-performance liquid chromatography (HPLC) for assay and degradation.[43,44] The addition of water allows the pH of the drug–excipient blend and the role of moisture to be investigated. Alternatively, thermal analysis methods, including differential scanning calorimetry (DSC) and isothermal calorimetry (ITC), are used to screen binary mixtures for their compatibility. However, data obtained via DSC/ITC are sometimes difficult to interpret or even misleading.[43,45]

A binary mixture approach is simple, but time and resource intensive, and may differ completely from a multicomponent system.[44] Due to these concerns, applicants can also perform drug–excipient compatibility studies via a prototype/trial formulation approach, with binary systems as a diagnostic back-up. In this complementary fashion, formulators may select one promising prototype formulation to move forward and identify any "culprit" excipient causing stability issues. The sponsor should use prior knowledge to justify inclusion or exclusion of certain excipients. For example, if the drug substance is a primary amine, and known to react with reducing sugar, the sponsor can include this as a rationale for not having lactose tested in the excipient compatibility studies. In addition, design of experiments (DOE) can be applied in the compatibility study, to determine the chemical interactions among components.

When the generic drug product is quantitatively and qualitatively the same as the RLD, in the absence of known stability problems, historical data may be acceptable in lieu of compatibility studies. However, applicants should be cautioned that excipients of different grades or excipients of the same grade, but sourced from different suppliers may contain different impurities that may interact with the drug substance. A systematic drug–excipient compatibility study will equip the product development scientist with a mechanistic understanding of formulation factors that affect drug product stability, and reduce formulation risk, leading to a decrease in development time and costs.

What is the rationale for the excipient selections?

Excipients can (1) aid in the processing of the dosage form during its manufacture; (2) protect, support, or enhance stability, bioavailability, or patient acceptability; (3) assist in product identification; or (4) enhance any other attribute of the overall safety, effectiveness, or delivery of the drug during storage or use.[46] Despite the fact that excipients can alter the stability, manufacturability, and bioavailability of drug products and are recognized as a major source of variability, the general principles of excipient selection are not well-defined, and excipients are often selected ad hoc. Note that different grades of excipients have different properties. For example, diluents may differ in particle size distribution and density and binders may differ in compressibility.

In this question, excipient choices are justified with respect to functions that are critical to product or process performance. It is important to indicate when pharmaceutical development work has demonstrated a functionality-related characteristic (ie, CMA, of an excipient plays a critical role for the manufacturability or quality attributes of the drug product). For example, when the polymer molecular weight and particle size are found to affect the release of a sustained release tablet, the relevant information should be documented in the answer. Note that compliance with USP-NF monographs is a minimum requirement. When necessary, stringent control of CMAs should be established beyond the compendial acceptance criteria to ensure consistent performance throughout the product lifecycle.[47] Formulators should anticipate lot-to-lot and supplier-to-supplier variability in CMAs and should establish specifications that will ensure consistent excipient performance.

P.2.2 Drug Product

What aspects of the formulation were identified as potentially high risk to the drug product performance?

Many attributes of the drug substance and excipients could potentially impact the CQAs of the drug product intermediates and finished drug product, making it unrealistic that a formulation scientist could investigate each identified material attribute of the drug substance and excipients during formulation development. Therefore, risk assessment leveraging common scientific knowledge, as well as the formulator's expertise, is valuable in prioritizing which material attributes warrant further study. Here, the applicant should provide a summary of the risk assessment approach used to rank or prioritize formulation variables based on their potential impact on drug product performance. It is up to the applicant to pick appropriately from a variety of recognized tools and to decide if the assessment should be formal or informal. The evaluation of the risk to quality should be based on scientific

knowledge and ultimately link to the protection of the patient; and the level of effort, formality, and documentation of the quality risk management process should be commensurate with the level of risk.[2]

Formulation variables that have been identified and justified as low risk are considered well understood. Those ranked medium risk can be considered acceptable based on current knowledge. Further investigation for medium-risk variables may be needed in order to reduce the risk to a lower level. Those ranked high risk are considered unacceptable and must be investigated to reduce the risk.

All potentially high-risk aspects of the formulation that could impact drug product performance should be addressed. For example, low drug loading of the final dosage form or key drug product intermediates may represent a high formulation risk due to the potential impact on content uniformity. Other formulation risks may include specific levels of excipients, as well excipient ratios. The level of rate controlling polymer in extended-release tablet formulations may impact drug release. The level of filler needed in a formulation could impact the tablet size. Specific excipient material attributes should also be considered in the overall formulation risk assessment and should tie into the rationale for excipient selection. For more complex dosage forms, an overall risk assessment of the formulation components (eg, IR granules, layered beads, coated beads, extra-granular excipients) can be performed followed by risk assessments of the variables involved to formulate each high-risk component (eg, for layered beads: bead selection, binder type and grade, solution viscosity, etc.). An acceptable response to this question summarizes the outcome of the risk assessment and provides the accompanying justification.

What formulation development studies were conducted? What attributes of the drug substance, excipients, and in-process materials were identified as critical and how do they impact the drug product CQAs?

By answering this question, applicants inform the FDA how they designed their product to meet the elements of the QTPP and to be safe and effective, in the case of innovator drugs, or those therapeutically equivalent to the RLD, in the case of generic drugs. A generic drug development scientist usually starts product design with studies of innovator product intellectual property, reverse engineering of the RLD, and preformulation. After studying the intellectual property of the RLD, the generic applicants may elect to go around the patent and certify that the patent is invalid, not infringed or unenforceable, or they may follow the patent and market the product after the patent exclusivity has expired.

Product understanding is exemplified by the ability to link input CMAs to output CQAs. The steps taken to gain product understanding may include the following[26]:

1. Identify all possible known input material attributes that could impact the performance of the product
2. Use risk assessment and scientific knowledge to identify potentially high-risk attributes
3. Establish levels or ranges of these potentially high-risk material attributes
4. Design and conduct experiments, using DOE when appropriate
5. Analyze the experimental data and, when possible, apply first principle models to determine if an attribute is critical
6. Develop a control strategy. For CMAs, define acceptable ranges. For material attributes shown to be noncritical, the acceptable range is the range investigated. When more than one excipient is involved, these defined acceptable ranges may be termed formulation design space

Formulation development studies are conducted to understand the impact of the potentially high-risk formulation variables identified in the preceding question on the quality of the finished drug product and to develop an appropriate control strategy to reduce the risk. For example, for problematic drugs, the formulation strategy may be to overcome limitations of the drug substance attributes such as poor solubility, poor dissolution, poor permeability, poor stability, or short plasma duration.

For each formulation development study used to identify which drug substance, excipient, or in-process material attributes (eg, particle size, blend uniformity, granulation flowability, etc.) are critical to manufacturing or product performance, a brief summary should be provided, including the objective of the study, the design and scale of the study, the factors that were investigated, the responses that were measured and how they compare with the pre-established targets, and the outcome of the study.

For complex formulations, the applicant usually investigates multiple candidate formulations during drug product development in order to obtain a bioequivalent product. This question encourages the applicant to share information on trial formulations or alternative mechanisms with the FDA. For instance, rapid dissolution of an immediate-release tablet may be achieved through particle size reduction or use of surfactant in the formulation. A controlled release product may be formulated with both immediate- and extended-release pellets or rather could employ a slow-release functional coating. Failed experiments often lead to better understanding of how formulation changes can alter performance and can help formulators identify the critical formulation or process variables. This development information may be useful to assess post-approval formulation changes, to develop

discriminating or predictive dissolution methods, and to set acceptance criteria on CMAs of the drug substance or excipients.

Applicants should also discuss how the final formulation was optimized. An optimization study is distinguished from an early screening study that explores a large range of space. An optimization study is more like fine tuning, and evaluates relatively small variations around a target formulation. Formulation optimization studies are essential to developing a robust formulation that is not on the edge of failure. Recent examples[48,49] have illustrated how formulations that are not optimized are risky because it is unknown whether any changes in the formulation itself or in the raw material properties would significantly impact the quality and performance of the drug product. However, it is not the number of optimization studies conducted, but rather the relevance and utility of the knowledge gained that is paramount to designing a quality drug product. As such, the QbD does not equal DOE, but the latter could be a valuable tool used in a QbD approach to development.[26]

How does the proposed commercial formulation differ from the formulations used during bioequivalence and/or clinical studies? What is the rationale for the formulation change? What biopharmaceutics evaluations (comparative dissolution, bioequivalence studies, biowaivers, etc.) support the formulation changes and link the development formulations to the proposed commercial formulation?

Applicants should summarize any planned formulation or finished dosage form physical attribute changes (eg, score configuration, shape, em/debossing, size) between the registration batches used for bioequivalence and/or clinical studies and the commercial batches. The rationale for such changes should be clear. The biopharmaceutical evaluation (eg, comparative dissolution, bioequivalence studies, biowaivers, etc.) conducted to support the changes should be fully described and its relevance scientifically justified. The evaluation should demonstrate that the proposed commercial formulation will not adversely impact the drug product quality, safety, and/or efficacy.

P.2.3 Manufacturing Process Development

A process is generally considered well-understood when: (1) all critical sources of variability are identified and explained, (2) variability is managed by the process, and (3) product quality attributes can be accurately and reliably predicted.[50] Similar to the steps outlined previously for product understanding, process understanding can be achieved by[26]:

1. Identifying all possible known process parameters that could impact the performance of the process
2. Using risk assessment and scientific knowledge to identify potentially high-risk parameters
3. Establishing levels or ranges of these potentially high-risk parameters
4. Design and conduct experiments, using DOE when appropriate
5. Analyze the experimental data and, when possible, determine scalability and apply first principle models to determine if a process parameter is critical. Link CMAs and critical process parameters (CPPs) to CQAs when possible
6. Develop a control strategy. For CPPs, define acceptable ranges. For process parameters shown to be noncritical, the acceptable range is the range investigated. When more than one process parameter or material attribute is involved, these defined acceptable ranges can be used to guide design space development, provided the ranges are identified via multivariate studies

The QbR questions next are designed to lead the applicant through these steps.

What is the rationale for selecting this manufacturing process for the drug product?

A pharmaceutical manufacturing process usually consists of a series of unit operations executed in batch mode or in a continuous manufacturing process. A unit operation is a discrete activity that involves physical or chemical changes, such as mixing, milling, granulation, drying, compression, and coating. The applicant should provide a schematic showing each unit operation in the drug product manufacturing process sequentially, including any reprocessing steps. The schematic should indicate all points of material entry for both raw materials and drug product intermediates and should include all in-process controls. For NDAs, the applicant can either provide a complete description of the commercial-scale manufacturing process or submit the proposed master production batch record. For ANDAs a detailed commercial-scale process description is also encouraged, but may not be submitted in lieu of the proposed or actual master production batch record.

For each unit operation, the process description should be supported by development work and should detail the following:

- batch size
- incoming raw materials (eg, excipients and processing aids and their grade, drug product intermediates from upstream unit operations) and the acceptance criteria for any material attributes known to be critical
- in-process controls and acceptance criteria
- equipment type, vendor, and model number
- process parameter settings (target and range)

Factors considered during manufacturing process selection may include the properties of the drug

TABLE 40.2 Examples of Pros and Cons of Typical Tablet Manufacturing Process[40]

Tablet manufacturing process	Pros	Cons
Direct compression	Simplified process, retains compactibility of materials	Segregation, flow
Dry granulation	Overcomes poor physical properties of API (particle size, shape)	Longer processing time, may compromise compactibility
Wet granulation	Improves uniformity, flow, and compactibility	Physical and chemical stability, residual solvents (nonaqueous granulation)

TABLE 40.3 Effect of Manufacturing Processes on Quality Attributes of the Drug Product

	Pre-blending and lubrication	Roller compaction	Milling	Final blending and lubrication	Compression
Assay	Medium	Low	Medium	Low	Medium
Content uniformity	High	High	High	Low	High
Dissolution	Medium	High	Medium	High	High
Purity	Low	Low	Low	Low	Low

substance, the desired characteristics of the drug product, and the complexity and robustness of the process. When applicable, the applicant should discuss in-house expertise with certain platform technologies or potential limitations, such as patent issues or facility and equipment availability to the extent that these scenarios guided manufacturing process selection.

Table 40.2 lists common tablet manufacturing processes and their pros and cons.[40] For example, for a drug substance with poor flowability and high drug loading, the granulation process may be considered to overcome the potential processing issues associated with direct compression. Once the process is selected, an applicant should focus on particular steps for which there are alternatives available (eg, selection of high-shear granulation vs fluid-bed granulation) with an explanation for the motivation behind each choice. Rationales can include scalability, yield, formulation aspects, expected granule characteristics, etc.[51] For granules that need to dissolve quickly, granulation techniques with lower shear forces (eg, fluid-bed granulation through top spray or bottom spray) can be selected. Fluid-bed granulation produces granules that have a more open, porous structure but are mechanically less stable.

Depending upon the product being developed, type of process, and process knowledge the development scientists have, it may be necessary to conduct preliminary feasibility studies prior to process development. Although it is not necessary to investigate alternative unit operations, if such studies were conducted to support a process choice, this information could provide very valuable support toward a better understanding of the final unit operations chosen to manufacture the drug product.

What is the potential risk of each process step to impact the drug product CQAs and how is the risk level justified?

This question is designed to connect each unit operation to its impact on the drug product CQAs and identify which process steps are high risk through risk assessment and prior knowledge. It is suggested that the applicant present the relationship between process steps and quality attributes of the drug product in the form of a matrix, as shown in Table 40.3.

Applicants should provide a concise summary of the risk assessment outcome, communicating the risk level assigned to each process step and clear justification for that assignment, whether the risk is high, medium, or low. The applicant should adequately justify low-risk process steps because these steps will not be studied extensively in the downstream development work. When justifying risk levels, the potential failure modes for each process step should be explained. For instance, if a process step is not properly controlled, the applicant should address which CQA of the drug product is likely to be affected and how it will be affected. For example, if the amount of a functional enteric coating applied to a tablet is not properly optimized, variable drug release will be observed, which may lead to noncompliance with the dissolution specification and ultimately a lack of efficacy in the patient.

For each of the potentially high-risk manufacturing unit operations:

a. *What input material attributes and process parameters were selected for study and what are the justifications for the selection?*

The previous question asks for the results of an overall risk assessment of the manufacturing process, which leads to the identification of high-risk steps that may impact the CQAs of the final drug product. For each high-risk unit operation, it is possible to identify CQAs of the drug product intermediate that are outputs of the unit operation and have a potential to impact finished drug product CQAs. Next, the process variables and material attributes, which may impact these CQAs of the drug product intermediate and, ultimately, the drug product CQAs, can be listed. Part (a) of this question focuses on the identification of

potentially high-risk material attributes and process parameters related to the identified high-risk steps that need to be investigated during development. Process parameters refer to the input operating parameters (eg, speed and flow rate) or process state variables (eg, temperature and pressure) of a process step or unit operation. Table 40.4 lists process parameters and input material attributes that could potentially impact the output quality attributes for typical pharmaceutical unit operations.[26] Equipment maintenance, operator experience and training, and standard operating procedures related to equipment use, testing, and facility supporting systems may link to product quality, directly or indirectly. Risk assessment allows the list of process parameters and material attributes to be prioritized to the vital few that pose the highest risk to impact the CQAs in order to focus development efforts. These variables are then investigated through controlled variations (eg, multivariate DOE and/or models) in order to better understand the process step and to develop a control strategy to consistently manufacture desired-quality drug product. This approach is illustrated schematically in Fig. 40.3, and can be used for each unit operation risk assessment.

The thought process depicted in the schematic is applied in the following example. Content uniformity is an element of the QTPP and is identified as a CQA for a low-dose, immediate-release tablet manufactured by blending and direct compression. The overall risk assessment of the manufacturing process identified the blending step to be high risk due to its potential impact on the content uniformity of the tablet. Next, blend uniformity is identified as a CQA of the blend (ie, drug product intermediate) because of its impact on the final tablet content uniformity. Drug substance and excipient particle size distribution and bulk density are material attributes and the number of blender revolutions and blender load level are process parameters that are most likely to impact the blend uniformity achieved during blending and are selected for further study during development.

If prior knowledge is used as the justification to fix a parameter or establish a starting point for investigation, then less justification is expected since experimental work will ensue. However, if prior knowledge is used in lieu of development studies, then more justification is expected in the application. The source and context of the prior knowledge, as well as its relevance to the current process, should be clearly communicated by the applicant.

b. *What process development studies were conducted? Provide a summary table listing batch size, process parameter ranges, equipment type, and estimated use of capacity.*

c. *What process parameters and material attributes were identified as critical and how do they impact the drug product CQAs?*

Parts (b) and (c) of this question are closely related. A concise summary of the conducted development studies that follow the logical progression of the drug product development should be provided. For each study, the objective of the study, the design and scale of the study, the factors that were investigated, the responses that were measured and how they compare with the pre-established targets, and the outcome of the study should be part of the answer. The summary should also include batch size, equipment type, equipment capacity used (or running time for continual operating equipment), process parameter set points, and ranges. A process parameter is considered critical when its intended variability is shown to have, or is likely to have, an impact on a CQA that is practically significant and, therefore, should be monitored or controlled to ensure the process produces the desired quality.

Criticality can be thought of as a continuum rather than discrete buckets (ie, critical or noncritical). When the underlying assumptions of the experiments used to determine criticality are changed, the conclusion on criticality can also change. Therefore, if future adjustments to process parameters and material attributes go beyond the investigated ranges discussed in the submission, then the conclusion on criticality should be reassessed.

At different stages of development, different approaches may be acceptable, such as screening, characterization, optimization, or verification studies. Process robustness is the ability of a process to consistently deliver acceptable drug product quality and performance while tolerating variability in the process and material inputs.[52] The effects of variations in process parameters and material attributes are investigated in process robustness studies. The relationship between input CMAs and CPPs and output CQAs is shown in Fig. 40.4. Studies to identify CMAs, CPPs, and their functional relationships to CQAs can be conducted at pilot- or laboratory-scale, and do not need to be conducted under cGMP. It is the responsibility of the applicant to demonstrate that the process is well understood, including how the interactions between the process parameters and/or material attributes impact the CQAs of the output materials, and to

TABLE 40.4 Typical Input Material Attributes and Process Parameters That May Impact Output Quality Attributes for Typical Pharmaceutical Unit Operations

Pharmaceutical unit operation		
Input material attributes	**Process parameters**	**Quality attributes**
BLENDING/MIXING		
• Particle size and distribution • Fines/oversize • Particle shape • Bulk/tapped/true density • Cohesive/adhesive properties • Electrostatic properties • Moisture content	• Type and geometry of mixer • Mixer load level • Order of addition • Number of revolutions (time and speed) • Agitating bar (on/off) • Discharge method • Holding time • Environment temperature and RH	• Blend uniformity • Potency • Particle size and distribution • Bulk/tapped/true density • Moisture content • Flowability • Cohesive/adhesive properties • Powder segregation tendency • Electrostatic properties
SIZE REDUCTION/COMMINUTION		
• Particle/granule size and distribution • Fines • Particle/granule shape • Bulk/tapped/true density • Adhesive properties • Electrostatic properties • Hardness/plasticity • Viscoelasticity • Brittleness • Elasticity • Solid state form/polymorph • Moisture content • Granule porosity/density *Ribbon milling* • Ribbon dimensions • Ribbon density • Ribbon porosity/solid fraction	*Impact/cutting/screening mills* • Mill type • Speed • Blade configuration, type, orientation • Screen size and type • Feeding rate *Fluid energy mill* • Number of grinding nozzles • Feed rate • Nozzle pressure • Classifier *Granule/ribbon milling* • Mill type • Speed • Blade configuration, type, orientation • Screen size and type • Feeding rate	• Particle/granule size and distribution • Particle/granule shape • Particle/granule shape factor (eg, aspect ratio) • Particle/granule density/porosity • Bulk/tapped/true density • Flowability • API polymorphic form • API crystalline morphology • Cohesive/adhesive properties • Electrostatic properties • Hardness/plasticity • Viscoelasticity • Brittleness • Elasticity
WET GRANULATION		
• Particle size distribution • Fines/oversize • Particle shape • Bulk/tapped/true density • Cohesive/adhesive properties • Electrostatic properties • Hardness/plasticity • Viscoelasticity • Brittleness • Elasticity • Solid state form/polymorph • Moisture content	*High/low-shear granulation* • Type of granulator (high/low shear, top/bottom drive) • Fill level • Pregranulation mix time • Granulating liquid or solvent quantity • Impeller speed, tip speed, configuration, location, power consumption/torque • Chopper speed, configuration, location, power consumption • Spray nozzle type and location • Method of binder excipient addition (dry/wet) • Method of granulating liquid addition (spray or pump) • Granulating liquid temperature • Granulating liquid addition rate and time • Wet massing time (post-granulation mix time) • Bowl temperature(jacket temperature) • Product temperature • Post mixing time • Pump Type: Peristaltic, Gear type • Granulating liquid vessel (eg, pressurized, heated) *Fluid bed granulation* • Type of fluid bed • Inlet air distribution plate • Spray nozzle (tip size, type/quantity/ pattern/configuration/ position) • Fill level	• Endpoint measurement (eg, power consumption, torque, etc.) • Blend uniformity • Potency • Flowability • Moisture content • Granule size and distribution • Granule strength and uniformity • Bulk/tapped/true density • API polymorphic form • Cohesive/adhesive properties • Electrostatic properties • Granule brittleness • Granule elasticity • Solid state form/polymorph

(*Continued*)

TABLE 40.4 (Continued)

Pharmaceutical unit operation		
Input material attributes	**Process parameters**	**Quality attributes**
	• Bottom screen size and type • Preheating temperature/time • Method of binder excipient addition (dry/wet) • Granulating liquid temperature • Granulating liquid quantity • Granulating liquid concentration/viscosity • Granulating liquid holding time • Granulating liquid delivery method • Granulating liquid spray rate • Inlet air, volume, temperature, dew point • Atomization air pressure • Product and filter pressure differentials • Product temperature • Exhaust air temperature, flow • Filter shaking interval and duration • Filter type and pore size	
DRYING		
• Particle size and distribution • Fines/oversize • Particle shape • Cohesive/adhesive properties • Electrostatic properties • Hardness/plasticity • Viscoelasticity • Brittleness • Elasticity • Solid state form/ polymorph • Moisture content	*Fluidized bed* • Inlet air volume, temperature, dew point • Product temperature • Exhaust air temperature, flow • Filter type and pore size • Filter shaking interval and duration • Total drying time *Tray* • Type of tray dryer • Bed thickness/tray depth (depth of product per tray) • Type of drying tray liner (eg, paper, plastic, synthetic fiber, etc.) • Quantity carts and trays per chamber • Quantity of product per tray • Drying time and temperature • Air flow • Inlet dew point *Vacuum/microwave* • Jacket temperature • Condenser temperature • Impeller speed • Bleed air volume • Vacuum pressure • Microwave power • Electric field • Energy supplied • Product temperature • Bowl and lid temperature • Total drying time	• Granule size and distribution • Granule strength, uniformity • Flowability • Bulk/tapped/true density • Moisture content • Residual solvents • API polymorphic form or transition • Purity profile • Moisture profile (eg, product temperature vs LOD) • Potency • Cohesive/adhesive properties • Electrostatic properties
DRY GRANULATION (ROLLER COMPACTION/CHILSONATION)		
• Particle size and distribution • Fines/oversize • Particle shape • Cohesive/adhesive properties • Electrostatic properties • Hardness/plasticity • Bulk/tapped/true density • Viscoelasticity • Brittleness • Elasticity • Solid state form/ polymorph	• Type of roller compactor • Auger (feed screw) type/design (horizontal, vertical or angular) • Deaeration (eg, vacuum) • Auger (feed screw) speed • Roll shape (cylindrical or interlocking). • Roll surface design (smooth, knurled, serrated, or pocketed) • Roll gap width (eg, flexible or fixed) • Roll speed • Roll pressure • Roller temperature • Fines recycled (yes or no, # of cycles)	• Ribbon appearance (edge attrition, splitting, lamination, color, etc.) • Ribbon thickness • Ribbon density (eg, envelop density) • Ribbon porosity/solid fraction • Ribbon tensile strength/breaking force • Throughput rate • API polymorphic form and transition

(*Continued*)

TABLE 40.4 (Continued)

Pharmaceutical unit operation		
Input material attributes	**Process parameters**	**Quality attributes**
EXTRUSION-SPHERONIZATION		
• Particle size, distribution • Fines/oversize • Particle shape • Cohesive/adhesive properties • Electrostatic properties • Hardness/plasticity • Bulk/tapped/true density • Viscoelasticity • Brittleness • Elasticity • Solid state form/ polymorph	• Type of extruder (screw or basket) • Screw length, pitch, and diameter • Screw channel depth • Screw blade configuration • Number of screws (single/dual) • Die or screen configuration (eg, radial or axial) • Die length/diameter ratio • Roll diameter (mm) • Screen opening diameter (mm) • Screw speed (rpm) • Feeding rate (g/min) • Type and scale of spheronizer • Spheronizer load level • Plate geometry and speed • Plate groove design (spacing and pattern) • Air flow • Residence time	*Extrudate* • Density • Length/thickness/diameter • Moisture content • API polymorphic form and transition • Content uniformity • Throughput *Pellets after spheronization* • Pellet size and distribution • Pellet shape factor (eg, aspect ratio) • Bulk/tapped density • Flowability • Brittleness • Elasticity • Mechanical strength • Friability
HOT MELT EXTRUSION		
• Particle size and distribution • Fines/oversize • Particle shape • Melting point • Density • Solid state form/ polymorph • Moisture content	• Screw design (twin/single) • Screw speed • Screw opening diameter • Solid and liquid feed rates • Feeder type/design • Feed rate • Number of zones • Zone temperatures • Chilling rate	• Extrudate density • Length/thickness/diameter • Polymorphic form and transition • Content uniformity • Throughput
TABLETING		
• Particle/granule size and distribution • Fines/oversize • Particle/granule shape • Cohesive/adhesive properties • Electrostatic properties • Hardness/plasticity • Bulk/tapped/true density • Viscoelasticity • Brittleness • Elasticity • Solid state form/ polymorph • Moisture content	• Type of press (model, geometry, number of stations) • Hopper design, height, angle, vibration • Feeder mechanism (gravity/forced feed, shape of wheels, direction of rotation, number of bars) • Feed frame type and speed • Feeder fill depth • Tooling design (eg, dimension, score configuration, quality of the metal) • Maximum punch load • Press speed/dwell time • Pre-compression force • Main compression force • Punch penetration depth • Ejection force • Dwell time	• Tablet appearance • Tablet weight • Weight uniformity • Content uniformity • Hardness/tablet breaking force/tensile strength • Thickness/dimensions • Tablet porosity/density/solid fraction • Friability • Tablet defects • Moisture content • Disintegration • Dissolution
ENCAPSULATION		
• Particle/granule size and distribution • Fines/oversize • Particle/granule shape • Cohesive/adhesive properties • Electrostatic properties • Hardness/plasticity • Bulk/tapped/true density	• Machine type • Machine fill speed • Tamping force • Number of tamps • Auger screw design/speed • Powder bed height	• Capsule appearance • Capsule weight • Weight uniformity • Content uniformity • Moisture content • Slug tensile strength • Disintegration • Dissolution

(*Continued*)

TABLE 40.4 (Continued)

Pharmaceutical unit operation		
Input material attributes	**Process parameters**	**Quality attributes**
• Viscoelasticity • Brittleness • Elasticity • Solid state form/ polymorph • Moisture content		
PAN COATING		
• Tablet dimensions • Tablet defects • Hardness/plasticity • Density • Porosity • Moisture content	• Type of pan coater (conventional or side-vented) • Pan (fully perforated or partial perforated) • Baffle (design, number, location) • Pan load level • Pan rotation speed • Spray nozzle (type, quantity, pattern, configuration, spray pattern) • Nozzle to bed distance • Distance between nozzles • Nozzle orientation • Total preheating time • Inlet air flow rate, volume, temperature, dew point • Product temperature • Individual nozzle spray rate • Total spray rate • Atomization air pressure • Pattern air pressure • Exhaust air temperature, air flow • Total coating, curing time and drying time	• Coating efficiency • Core tablet weight before and after preheating • Moisture (gain/loss) during preheating • Environmental equivalency factor • Coated drug product (eg, tablet or capsule) appearance • Percentage weight gain • Film thickness • Coating (polymer and/or color) uniformity • Hardness/breaking force/tensile strength • Friability • Moisture (gain/loss) during overall process • Residual solvent(s) • Disintegration • Dissolution • Tablet defects • Visual attributes
FLUID-BED COATING		
• Tablet dimensions • Tablet defects • Hardness/plasticity • Density/porosity • Moisture content	• Type of fluid-bed coater • Fluid-bed load level • Partition column diameter • Partition column height • Number of partition columns • Air distribution plate type and size • Filter type and pore size • Filter differential pressure • Filter shaking interval and duration • Spray nozzle (type, quantity, pattern, configuration) • Nozzle port size • Total preheating time • Spray rate per nozzle • Total spray rate • Atomization air pressure • Inlet air flow rate, volume, temperature, dew point • Product temperature • Exhaust air temperature, air flow • Total coating, curing and drying time	• Coating efficiency • Core tablet weight before and after preheating • Moisture (gain/loss) during preheating • Environmental equivalency factor • Coated drug product (eg, tablet or capsule) appearance • Percentage weight gain • Film thickness • Coating (polymer and/or color) uniformity • Hardness/breaking force/tensile strength • Friability • Moisture (gain/loss) during overall process • Residual solvent(s) • Disintegration • Dissolution • Tablet defects • Visual attributes
LASER DRILLING		
• Tablet size/dimensions • Polymer type • Membrane thickness	• Conveyor type • Conveyor speed • Laser power • Number of pulses • Type(s) of lens(es) • One or two sided • Number of holes	• Opening diameter (internal and external) • Depth • Shape of the opening

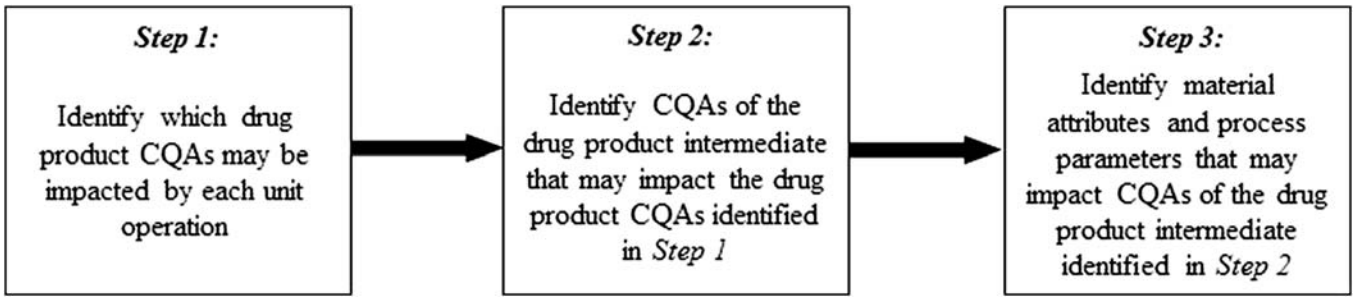


FIGURE 40.3 Schematic of the method used to identify process variables for further study.

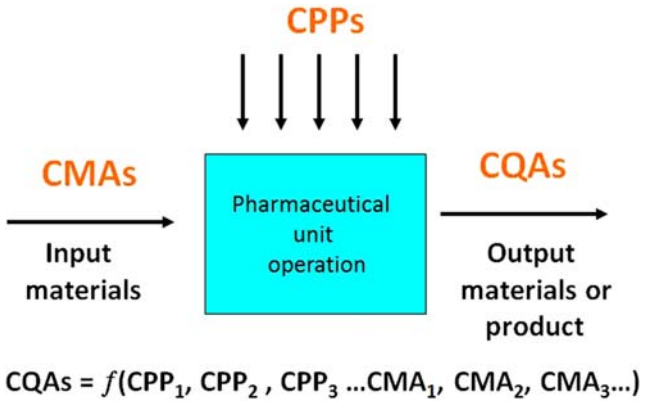


FIGURE 40.4 Linking input critical material attributes (CMAs) and critical process parameters (CPPs) to output critical quality attributes (CQAs) for a pharmaceutical unit operation.

establish limits for these CPPs (and CMAs) within which the quality of drug product is assured.

When a DOE study is performed, the *alpha* level used to judge whether a test statistic is significant should be provided and justified if it differs from the commonly used *alpha* level of 0.05. It is important to ensure that data are analyzed appropriately so that the conclusion is statistically meaningful. For example, analysis of variance (ANOVA) tables tell the overall story of the model, investigating and modeling the relationship between a response variable and one or more independent variables. When center points are included, curvature effect can be assessed and included in the ANOVA table. The saying that a picture is worth a thousand words really holds true when discussing DOE results. Pareto charts and half-normal plots are helpful when talking about what factors are significant. Main effect plots and interaction plots are also good visual aids. Contour plots are useful to see how two factors jointly affect responses and overlay plots support the selection of process parameter ranges.

When a trial and error or one-factor-at-a-time approach is taken, the applicant needs to explain how factors were determined to be significant, as well as how decisions regarding criticality were made. Information about interactions cannot be gained using this approach. Therefore, the applicant has a higher burden when justifying the chosen proven acceptable ranges.

Based on the outcome of the process development studies, the applicant can identify the process parameters, material attributes, and any interactions determined to be critical and can link these variables back to the drug product CQAs. This discussion of CPPs and CMAs conveys the process understanding gained throughout development and lays the foundation for proposing a control strategy. If no material attributes or process parameters are found to be critical, the range investigated is the range that should be reflected in the control strategy.

d. *How were the process parameters adjusted across lab, pilot/registration, and commercial scale? What are the justifications for any changes?*

Most manufacturing processes, whether as simple as making an aqueous solution of highly soluble drug substance or as complex as coating extended-release pellets to be used in a multiparticulate tablet formulation, have certain scale-dependent process parameters that will require adjustments when the manufacturing process is transferred to larger scale. Therefore, additional experimental work may be needed. Prior knowledge can play a very significant role in this regard because most pharmaceutical companies tend to use the same manufacturing technologies and raw materials on a regular basis. The functional relationship between CMAs, CPPs, and CQAs is also considered during scale-up. Monitoring CQAs of in-process material makes scaling less equipment-dependent as opposed to only monitoring CPPs.

Under part (d) of this question, the applicant should discuss the scale-up plan that builds upon experience obtained from development studies and/or the production of exhibit batch(es), scientific theory, first principles, literature references, vendor scale-up factors, laboratory scale to exhibit scale process transfer for this product, exhibit batch production, dimensional analysis, and/or mechanistic modeling. For reference to other applications, similarity of dosage form, physical properties of the drug and drug loading, excipients, equipment, and change of scale should be considered. Pharmaceutical development scientists can also make use of computer-aided process design (CAPD), and process simulation, such as roller compaction simulation, and tablet simulation, to support process development and optimization of manufacturing.[53] The increased use of CAPD and process simulation in the pharmaceutical industry should promise more robust processes are developed faster, and at a lower cost, resulting in higher quality products.

Changes made to the process parameters can be summarized in a table. An example of scale-up from lab-scale to pilot-scale and then to commercial-scale based on Johanson's model[54] for a roller compaction unit operation is summarized in Table 40.5.

The volume occupied by the batch in the processing equipment at each scale should be provided along with an assessment of capacity utilization differences on product quality and manufacturability. This assessment should also include the minimum and maximum recommended capacity utilization of the corresponding equipment at commercial scale. A similar comparison for the run time of continuous manufacturing equipment should also be provided to evaluate if/when a state of control is reached.

Equipment with different geometric design and operating principles (eg, bin blenders vs ribbon blenders) can impact blend homogeneity. In addition, various tumble blenders have significantly different designs and, therefore, operating efficiencies. For example, an asymmetrical blender such as a slant cone can be highly efficient in mixing when compared with a symmetrical blender such as a V-blender. A correlation between the two types of tumble blenders should be established through experimental studies if a change in equipment is anticipated during scale-up.

It is at the applicant's discretion to submit scale-up data, such as actual commercial-scale process performance qualification information in the application for a complex drug product that has a high risk of scale-up failure; however, in some cases it may be requested by the FDA prior to approval.

If applicable, what online/at line/in-line monitoring technologies are proposed for routine commercial production that allow for real-time process monitoring and control? Provide a summary of how each technology was developed.

Design space may be scale and equipment dependent. A limitation of design space based solely on process parameters is that the verification burden at different scales or in different equipment can limit the sought-after flexibility. An alternative to a control strategy that hinges solely on process parameter-based design space is a control strategy that actively monitors the output CQAs of intermediates (eg, blend uniformity) using process analytical technology (PAT) to adaptively adjust the input variables through feedback/feed-forward mechanisms (eg, blending time).[40]

Application of PAT involves four key components[26,55]:

- Multivariate data acquisition and analysis, which integrates product and process understanding into process control to achieve CQAs
- Process analytical chemistry tools that provide real-time and in situ data about the process status
- Process monitoring and control
- Continuous process optimization and knowledge management

TABLE 40.5 Scale-Up of the A Roller Compaction and Integrated Milling Process

Scale	Batch size (kg)	Batch size (units)	Alexanderwerk model	Roller width (mm)	Roller diameter (mm)	Roller gap (mm)	Roller pressure (bar)	Mill screen orifice size (mm)
Lab	5.0	25,000	WP120	25	120	1.2–2.4	20–77	1.0
Pilot	50.0	250,000	WP120	40	120	1.8	50	1.0
Commercial	150.0	750,000	WP200	75	200	2.0–4.0[a]	31–121[a]	1.0

[a]*The range is based on the scale-up equation and needs to be verified.*

TABLE 40.6 Process Analytical Technology (PAT) Implementation in Solid Oral Unit Operation and Its Benefits in Product Development[56]

Unit operation	PAT technologies employed	Benefits in product development
Particle size reduction	NIR spectroscopy	Design/control milling process of API and in-process agglomerates
Blending/ mixing	NIR spectroscopy FT-Raman spectroscopy	Identify safe-zone and the most appropriate mixing time
Granulation	NIR, Raman XPRD spectroscopy Thermal effusivity Combination of an image-process device and a fuzzy control system Acoustic emission Stress fluctuation spectroscopy	Justify granulation process parameter ranges with controlled granulation quality
Drying	NIR spectroscopy	Identifying drying curve and end point Justify moisture content specifications
Coating	NIR spectroscopy Raman spectroscopy	Determine coating rate and end point with minimized manufacturing variability; interpret coating uniformity
Unit dosing (compression or encapsulation)	NIR spectroscopy Raman spectroscopy	Detect drug distribution uniformity in finished product. Identify undesired process deviation

PAT can be utilized offline, at line, online, and in line, and its implementation in different solid dose unit operations is summarized in Table 40.6.[57] For instance, in the blending process, the use of online near-infrared (NIR) can evaluate the mixing status of active pharmaceutical ingredient (API) and excipients and monitor compositional variability over the entire blending time, rather than the traditional one-time sampling points at the end of blending. This allows for the evaluation of segregation or demixing potential, and identifies the safe-zone of uniform mixing or the most appropriate mixing time. Understanding of critical processes, as well as justification of process parameters and in-process specification ranges via PAT work, provides superior control quality. Ideally, PAT can help provide feedback to the control system, and manufacturers have the ability to effect continuous improvement and continuous real-time assurance of quality. Nevertheless, there are some regulatory, quality assurance, cGMP, and validation challenges associated with the adoption of PAT. Pharmaceutical industries are strongly encouraged to prepare for the challenges, and promote drug product development and process understanding via PAT efforts. The FDA has approved a number of applications that implemented PAT, including drug substance and drug product manufacturing processes for innovator products, generic products, and veterinary products, demonstrating success with this initiative.

In response to this QbR question, a description of the online/at line/in-line technology that is being used for real-time process monitoring and control should be provided as well as a discussion of how the technology will be implemented and the impact of the technology on the overall control strategy. A summary of the method development would suffice if the technology is being used for process monitoring and control. However, implementing PAT for real-time release testing (RTRT) of the drug product would require additional supportive information to be provided. The applicant should provide a description of the instrument and its location, information on the development and validation of the calibration method, and a summary of the model maintenance approach. Another consideration when PAT is utilized is having a contingency plan detailing how the process will be controlled in the event that the PAT analyzer fails in the middle of manufacturing or is not available from the very beginning of batch manufacturing.

P.2.4 Container Closure System

What specific container closure system attributes are necessary to ensure drug product integrity and performance through the intended shelf life? If applicable, what are the differences in the container closure system(s) between this product and the RLD?

The applicant may give the rationale (eg, drug product stability concerns) for selecting a particular container closure system including light resistance, moisture protection, or inert atmosphere. For example, when atmospheric oxygen is the source of drug oxidation, the formulator can select an appropriate packaging material (material with lower oxygen transmission rate), oxygen scavenger, and/or packaging method (packaging under inert gas) to overcome the oxidation problem.[42] Controlling the oxygen permeability and headspace oxygen in bottles is appropriate for drug product shelf stability, but cannot ensure drug stability after the consumer has opened the product. Therefore, blister packaging

may need to be used for some extremely oxygen-sensitive drugs.

Consideration should also be given to the intended use of the drug product and the suitability of the container closure system for shipping and handling. When applicable, the applicant should summarize studies illustrating that tablet breakage or delamination is minimized. Another consideration is the need for child protection using child-resistant closures. In some cases, data supporting that drug product quality is not adversely impacted by temperature excursions may be warranted.

When discussing differences in the container closure system between the proposed generic and the RLD for a solid oral dosage form, applicants should detail the material of construction, container size(s) and fill count/volume/weight, filler type, desiccant type and desiccant configuration (sachet, canister, etc.), and quantity.

How was the container closure system(s), including bulk containers, qualified for suitability (protection, compatibility, safety, and performance)?

There are four elements that need to be considered when qualifying a container closure system as suitable for its intended use: protection, compatibility, safety, and performance.[58] The system should adequately protect the dosage form to alleviate stability concerns. It should also be compatible with the dosage form. Possible interactions between the product and container or label components should be considered. The container closure system should also be composed of materials that are considered safe for use with the dosage form and the route of administration. Finally, if the packaging system has a performance feature in addition to containing the product, it should be shown to function properly and similarly to the RLD, as applicable. If the only rationale provided for qualifying a container closure system is that the same system was previously FDA-approved for a different drug product in a different application, this justification is generally inadequate to ensure the proposed container closure system is suitable for the current drug product under review.

P.2.5 Microbiological Attributes

When applicable, what microbiological attributes were evaluated on the finished product?

Solid oral dosage forms represent a relatively low microbiological risk to patients; therefore, the microbiological controls associated with their manufacture are not as stringent as those for sterile or aqueous nonsterile products. However, microbiological control is important in the manufacture of solid oral dosage forms in order to prevent drug product degradation, the introduction of microbial metabolites and/or toxins during the manufacturing process, and patient exposure to high numbers of harmful species of microorganisms. Minimally processed components derived from certain plant and animal sources, or manufacturing processes involving extended aqueous holding times, may carry a higher overall microbiological risk.

The rationale for performing or not performing microbial limits testing for nonsterile drug products should be evaluated (eg, Decision Tree #8 in ICH Q6A, USP <1111> and USP <1112>).[59–61] Solid dosage forms that have low water activity or are otherwise inherently antimicrobial are excellent candidates for reduced microbial testing for product release and stability.[59] A submission that proposes a waiver of microbial enumeration testing on product release should include a description of the following:

- Critical control points in the manufacturing process that may affect the bioburden of the finished drug product
- Monitoring schedule and acceptance criteria for the critical control points documented in the batch record per 21 CFR 211.188
- Microbiological testing results for stability or exhibit batches of the drug product
- Proposed stability testing schedule

Critical control points may include establishing bioburden limits on incoming materials, defining the maximum processing time for any aqueous processing steps, and justifying the maximum holding time for any aqueous coating solutions with supporting data. If extensive product history is available, and upstream controls are established, it may be feasible to propose a waiver of all microbial limits studies, including those on stability. When a significant history of product and process control is still being developed at the time of application submission, microbial enumeration testing at batch release should be proposed.

Due to the hygroscopic nature of gelatin and its greater water activity, gelatin may be capable of supporting the growth of some microorganisms over the product shelf-life, even if adequate microbiological manufacturing controls have been implemented. Stability batches of drug products containing gelatin should undergo microbial limits testing.

USP <1111> provides a risk-based standard for total aerobic microbial counts, total combined yeasts/molds count, and specified microorganisms in nonsterile drug products based upon the route of administration.

If alternate microbial limits acceptance criteria are proposed, the applicant should provide a justification for the change and supporting data. It is not required that applicants utilize USP <61> and <62> test methods for microbial enumeration. However, a detailed

description of the microbial limits and specified organism detection test methods, as well as verification of the methods suitability for the finished drug product, need to be provided.

P.2.6 Compatibility

If applicable, what supportive data demonstrates the compatibility of the drug product with the means of administration (eg, additives and/or diluents, other co-administered drugs, dosing device)?

For some solid oral drug products, further manipulation or preparation may be required prior to administration. In cases such as these, appropriate compatibility information should be submitted in the application to support the drug product labeling. For a generic drug product, assurance that the generic drug meets all of the performance characteristics of the RLD is needed. When evaluating the QTPP for the proposed drug product, thorough consideration of any alternative means of administration should be made early in the process. This is the best approach to ensure that the appropriate development studies have been conducted to support all aspects of compatibility.

Examples where supportive data to demonstrate compatibility are required include capsules containing beads that are labeled for "sprinkle over applesauce" or to be dissolved in a glass of water. Some tablets need to be crushed and suspended for administration to patients who have difficulty swallowing or for administration to a pediatric population. For instance, if the labeling indicates that "the suspension is homogenous and can be stored for either up to 30 days at room temperature (below 30°C/86°F) or up to 75 days at refrigerated conditions (28°C/35–46°F) in the glass bottle," then stability data demonstrating that the proposed drug product is stable when prepared and stored according to the conditions (eg, temperature and time) specified in the label should be provided. Furthermore, some tablets are intended for administration through a nasogastric tube or oral syringe. For these tablets, applicants should demonstrate that when the label instructions are followed, the tablet disperses into granules or particles that are small enough to ensure that the nasogastric tube or syringe used for administration is not blocked or clogged. Similarly, the same is true for capsules and bead size.

The previous examples are meant to be illustrative and not exhaustive. The data required varies on a case-by-case basis and is dependent on the means of administration and the specifics of the drug product and the associated labeling. All alternative routes of administration should be fully evaluated with respect to drug product compatibility.

P.3 Manufacture

Who manufactures the drug product? List each participant and facility involved in drug product manufacturing/testing activities and clearly state their function. List the date of the last FDA inspection of each facility involved and the result of the inspection. Has the manufacturer addressed all concerns raised at the FDA inspection?

It is imperative that the agency has a complete list of facilities associated with the application early in the submission process to ensure that the necessary facility evaluations can be performed and any necessary preapproval inspections scheduled to allow timely approval. Applicants should include the name, address, and FEI number for each site or facility involved with commercial manufacturing, packaging, or testing the drug product. Sites used to conduct stability testing of the registration batches supporting the submission should also be included if these sites differ from those planned to be used for commercial batches. When describing the responsibility of each testing site, it should be clear what component is being tested (eg, excipient, packaging component, drug substance, drug product intermediate, drug product), what test is being performed, and at what stage of manufacture (eg, in-process, release, stability). The dates of the last FDA inspection and the outcome of the inspection should be noted. If any concerns were raised by the FDA during the inspection, then a summary explaining how those concerns have been adequately addressed should be provided.

What is the commercial batch formula and how does it differ from the registration batch formula? Provide justifications for any differences.

A side-by-side comparison of the registration and commercial batch formulae should be provided. If there are any differences between the registration batch and the commercial batch formula, these should be noted and discussed in terms of the rationale for the differences as well as the potential impact on product manufacturability and CQAs. The weight or measure of each component of the batch formula should be listed (as per section 2.3.P.1.). It is helpful to also include each component on a weight percent basis. The name of the component should include a reference to its quality standard (eg, USP, in-house, etc.). For excipients where a range has been justified, the target amount and range should be included in the batch formula.

Drug substance overages are generally discouraged; however, if any an overage is used, it should be made clear and justification should be provided. Overages are added to the formulation either for a manufacturing loss that has been demonstrated to

be reproducible, or for stability concerns. The composition should reflect the actual quantities charged to the batch since the overage is a part of the formulation. For example, a product with 2% overage of the drug substance should list the drug substance as 102% in the formulation. Note that stability assessment should be based on the actual quantity used in the formulation (ie, 102%).

Processing agents (eg, water, solvents, and nitrogen or other gases) that do not remain in the finished product should be included. Any gases used during manufacture should be listed and their purpose identified (eg, blanket formulation, fill vial headspace). An asterisk after the ingredient name with a footnote below the table could be used to identify components that are removed during processing or to clarify the purpose of inert gases used during the manufacturing process.

An excess is any inactive ingredient prepared in addition to the overall formulation to compensate for certain losses during manufacturing, but that does not become part of the product formulation once manufacture is completed. For example, a certain percentage of excess coating solution is often prepared to compensate for losses of the coating solution to the walls of the equipment. Excess material can be denoted with an asterisk and explained in a table footnote. The excess material should be reasonable and well justified.

If a common formulation is used to produce multiple products' strengths, then the composition of each strength should be provided including the total weights/volumes as applicable. If the end product is a result of two or more sublots, then the composition of each sub-lot and the number of sublots should be clear. If the registration batch is manufactured as one lot, but the commercial scale batch will be manufactured using sublots, then this should also be indicated.

In some instances, separately blended or formulated materials that are later combined during manufacturing should be listed separately in the batch formula. For example, some modified-release products contain a mixture of immediate release and extended-release beads within a dosage unit. In this case, the composition table should present the batch formulas for the individual subcomponents of the dosage unit.

What is the flow diagram of the manufacturing process that shows all incoming materials, processing steps/unit operations, and in-process controls?

The flow diagram is a schematic showing each unit operation of the commercial manufacturing process consecutively from first charging of materials through packaging. It should also show each point of material entry. The material entry into a particular unit operation may be drug substance, raw materials, or drug product intermediates. The flow diagram should also include the in-process controls for each unit operation. In cases where reprocessing steps are possible, those should be included in the diagram.

What is the detailed process description, including process parameters, material attributes of raw materials and intermediates, equipment type, batch size, in-process controls, including acceptance criteria and any proposed reprocessing?

Applicants should submit a manufacturing process description for a representative commercial batch of the drug product that has the level of detail comparable to a master production batch record (NDAs) or a proposed or actual master production batch record (ANDAs). For each process step and proposed reprocessing step, the narrative should detail:

- Batch size
- Incoming raw materials/drug product intermediates and acceptance criteria for any CMAs. For raw materials, the grade should also be specified
- In-process controls and acceptance criteria
- Equipment description (type, vendor, model number, etc.)
- Process parameter settings (target and range), indicating which process parameters are critical

Process parameter ranges and in-process controls should be supported by pharmaceutical development work discussed in 3.2.P.2.3. A risk-based approach should be taken when establishing hold times for drug product intermediates. Data should be provided for studies conducted in the proposed storage container under the proposed storage conditions for the appropriate length of time to demonstrate that holding does not adversely affect quality.

What in-process sampling strategies and methods are used to monitor in-process material attributes that have a potential to affect quality?

Details about sampling for each unit operation (eg, sample size and/or quantity, sampling frequency, sampling location) should be provided for CQAs of drug product intermediates. For example, details about blend uniformity sampling in bins or drums or stratified sampling of core tablets should be provided for low-dose products that have a high segregation potential.

What are the in-process test results for each process step of the registration batch(es)? What are the differences, if any, in the in-process controls for the registration batch (es) and the intended commercial batches? What are the justifications for these differences?

It is best to present the in-process test results for each step of the process as a tabular summary listing the test, the acceptance criteria, and the observed results. Those quality attributes of the drug product intermediates that may have a significant impact on the quality of the finished drug product should be identified and monitored. In addition, if there are any

differences between the tests performed and/or the acceptance criteria between the registration batches and proposed commercial-scale batches, the differences should be explained and well justified. It should also be made clear if any of the tests are planned through validation, but not as part of the commercial manufacturing control strategy.

P.4 Control of Excipients

What are the excipient specifications and how are they justified? How do the proposed acceptance criteria for the material attributes of the excipients ensure the quality of the final drug product?

USP-NF excipients should minimally comply with the current respective monograph. In-house analytical procedures used to test compendial excipients should be appropriately validated. Any additional test(s) needed for CMAs identified during development beyond those listed in the monograph should also be included in the specification and the response to the QbR question should explain how the CMA and the proposed acceptance criterion relates to the final drug product quality. Applicants should discuss the adequacy of the specification to mitigate any associated risks to manufacturability or drug product performance for excipients where lot-to-lot variability is anticipated. Specifications including analytical procedures and method validations for the noncompendial excipients are also expected.

Applicants should provide both in-house and vendor full testing results for excipient lots used to manufacture the registration batch(es) or a representative CoA for any additional lot that has been fully tested. Note that proposals for vendor qualification and skip-lot testing are beyond the purview of the review divisions.

In cases where a novel excipient not previously used in a CDER-approved product is included in the formulation, information detailing the excipient manufacture, characterization, control, and safety qualification, or a reference to a DMF containing the previously mentioned information should be included. Applicants should provide bovine spongiform encephalopathy (BSE)/transmissible spongiform encephalopathy (TSE) certification as well as the country of origin for any excipient of animal origin.

P.5 Control of Drug Product

What is the drug product specification, what is the justification, and how is it linked to the product performance and patient safety? Does the specification include all the CQAs for the drug product?

How do the batch analysis results compare with your proposed specification? Provide a summary of the batch analysis results.

To illustrate related concepts, these questions are grouped together in this chapter, although they are presented differently in the QbR assessment system. A specification is a list of tests, references to analytical procedures, and appropriate acceptance criteria, which are numerical limits, ranges, or other criteria for the tests described.[59] The drug product should conform to the specification to be considered acceptable for its intended use. Specifications are quality standards that are proposed and justified by the applicant and approved by the FDA as conditions of approval. The specification is part of the total control strategy designed to ensure product quality and consistency. To answer this question, the applicant should present a tabular summary of the specification including columns for the test, analytical procedure, acceptance criteria, and representative results for at least one registration batch of each strength. The specification should focus on those characteristics that are useful in ensuring the safety and efficacy of the drug product including all drug product CQAs necessary to ensure the identity, strength, purity, and quality. For USP drug products, the specification should include all tests listed in the monograph as well as any additional tests that are deemed necessary for the drug product. The applicant should also provide justification for the chosen acceptance criteria for each CQA as well as a discussion linking the proposed control of the drug product to in vivo performance and patient safety. In cases where any tests of CQAs are omitted, scientific rationale should be provided.

In some cases, it may be appropriate to measure and control the drug product CQAs in-process in lieu of end-product testing. Controlling certain quality attributes upstream where the measured response provides a timely signal to make process adjustments is an effective strategy for ensuring product quality. For instance, if organic solvent is used during wet granulation, it is advantageous to control residual solvents at this step when drying time can be adjusted rather than waiting to test the compressed tablets when the only recourse for noncompliance with the residual solvent specification is batch failure. In addition, RTRT (formerly called parametric control) can be employed when the model is predictive of in vivo performance and captures the dominating factors that contribute to variability in the CQA. As a theoretical example, attributes relating to the properties of a tablet granule, such as porosity, particle size, surface area, or bulk/tapped density, if shown to have a predictive relationship with dissolution behavior, could serve as RTRT surrogates for dissolution testing.[62] Both of these approaches should be clearly presented and appropriately validated to show that test results or product

performance characteristics do not change from the in-process stage to the finished product.

The applicant should discuss how the registration batch results compare with the proposed specification. Whenever possible, results should be presented quantitatively rather than in general terms such as "complies" or "meets limit." Representative results for one registration batch of each strength can be presented with a hyperlink or reference to where additional batch results can be located in Module 3. In cases where the CQA still complies with the specification, but is on the edge of failure (eg, low assay, high Acceptance Value (AV) for content uniformity, high impurity level), the applicant may wish to discuss the root cause of the observed result and detail any further development or optimization that is planned to ensure that the CQAs will be achieved when the drug product is commercialized.

For each test in the specification, provide a summary of the analytical procedure(s) and, if applicable, a summary of the validation or verification report(s).

Applicants should identify regulatory and alternative analytical procedures when multiple analytical procedures can be used for a test. Note that the analytical procedures in the USP-NF are those legally recognized under section 501(b) of the FD&C Act (the Act) as the regulatory analytical procedures for compendial items and are the prevailing procedures used for the purposes of determining compliance with the Act. Applicants may propose to use a validated alternative analytical procedure found in other compendia, such as the BP, the EP, or the Japanese Pharmacopoeia (JP), instead of the regulatory analytical procedure.[63] The applicant should provide a rationale for its inclusion, as well as validation and comparative data to show that the alternative procedure performs equal to or better than the regulatory analytical procedure.

It is not necessary to provide a summary of the analytical procedure in the QOS when the USP procedure is followed, although the procedure details should still be submitted in Module 3. Each noncompendial procedure should be summarized in a tabular or descriptive form. It should include the critical parameters for the method and system suitability criteria, if applicable. For impurity methods where impurities are not quantified using impurity standards, the relative response factors should be provided.

Quantitative noncompendial analytical methods such as Karl Fischer, assay, impurities, and residual solvents should be fully validated per the recommendations found in ICH Q2 *Validation of Analytical Procedures: Text and Methodology*[64] and a summary table of the validation performance characteristics should be provided. The stability-indicating capability of assay and impurity methods should be demonstrated through HPLC peak purity and achievement of adequate mass balance. Stress studies should target 5−20% degradation to avoid secondary degradation. For very stable molecules that are difficult to degrade, a justification should be provided (ie, stress conditions that go beyond the usual) along with a summary of forced degradation results. Applicants should attempt to quantitate the degradants observed for each stress condition individually by name or relative retention time. The agency expects that any loss in the active peak is compensated by a corresponding increase in the amount of degradants. Any failures to obtain adequate mass balance should be fully explained.

According to USP <1226>, users of analytical methods described in USP−NF are not required to validate the accuracy and reliability of these methods, but merely verify their suitability under actual conditions of use.[65] Verification is not required for basic compendial test procedures including, but not limited to, loss on drying, residue on ignition, and pH that are routinely performed unless there is an indication that the compendial procedure is not appropriate for the article under test. A summary table of the verification performed for compendial methods should be provided. Typical performance characteristics include specificity and precision for assay methods and specificity, precision, linearity, and LOD/LOQ for impurity methods. In all cases, the applicant should demonstrate that the system suitability criteria can be achieved under actual conditions of use.

The response to this question may contain a reference or hyperlink to Module 3, the location of the full validation or verification information.

What are the drug product degradants? For each degradant, what is the structure, chemical name, origin, and mechanism of formation? How are the proposed limits justified and/or qualified for safety based on nonclinical studies? What is the control strategy for the potential drug product degradants?

Degradation products expected to occur during manufacture of the commercial product or under recommended storage conditions should be addressed in the proposed drug product specification. These degradants result from a chemical change in the drug substance brought about during manufacture and/or storage of the new drug product by the effect of, for example, light, temperature, pH, water, or by reaction with an excipient and/or the immediate container closure system.[66]

Generally, it is not necessary to test the drug product for quality attributes uniquely associated with the drug substance. Thus, process impurities present

in the drug substance need not be monitored or specified in the drug product unless they are also degradation products. If the applicant chooses to include drug substance process impurities or synthetic precursors in the drug product specification (eg, due to their presence in chromatograms or as markers in chromatograms), then those impurities should be limited at release and on stability to the same level as in the drug substance specification.

For each identified degradation product, the structure, chemical name(s), and origin should be provided. The applicant should also describe any studies performed to identify the degradant and its mechanism of formation, as well as the control strategy in place to minimize formation to the extent possible. When justifying the proposed acceptance criterion for the degradants, applicants should consider patient safety, particularly for degradants with the potential to be genotoxic or carcinogenic. In cases where applicants propose a degradant level above those recommended in the current guidance for industry for impurities in drug products and ICH Q3B,[66–68] the degradant will need to be qualified by conducting toxicity studies, submitting relevant scientific literature, demonstrating that the degradant is a significant human metabolite of the drug substance, or by showing that the RLD contains similar levels of the degradant using the same validated, stability-indicating, analytical procedure.

Degradants should be reported with the appropriate number of decimal places. Below 1.0%, the results should be reported to the number of decimal places in the applicable reporting threshold whereas at and above 1.0%, the results should be reported to one decimal place.

What is the proposed control strategy for the drug product manufactured at commercial scale? What are the residual risks upon implementation of the control strategy at commercial scale?

A control strategy is a planned set of controls derived from current product and process understanding that ensures process performance and product quality.[3,4] It can include, but is not limited to, input material attributes, facility operating conditions, and equipment operating conditions (ie, process parameters, in-process controls, specifications, sampling plans, and frequency of monitoring and control). Systematic and well-executed pharmaceutical development studies that result in enriched product and process understanding and identification of all CMAs and CPPs will culminate in the design of an appropriate control strategy to ensure the identity, strength, quality, and purity of the drug substance and the bioavailability and performance of the drug product throughout its lifecycle. Fig. 40.5 shows pictorially the different levels of control that are possible.

The most basic control is to have knowledge of the operating range, as well as the proven acceptable range of the process parameters. Proven acceptable range is defined as a characterized range of a process parameter for which operation within this range, while keeping other parameters constant, will result in producing a material meeting the relevant quality criteria.[4]

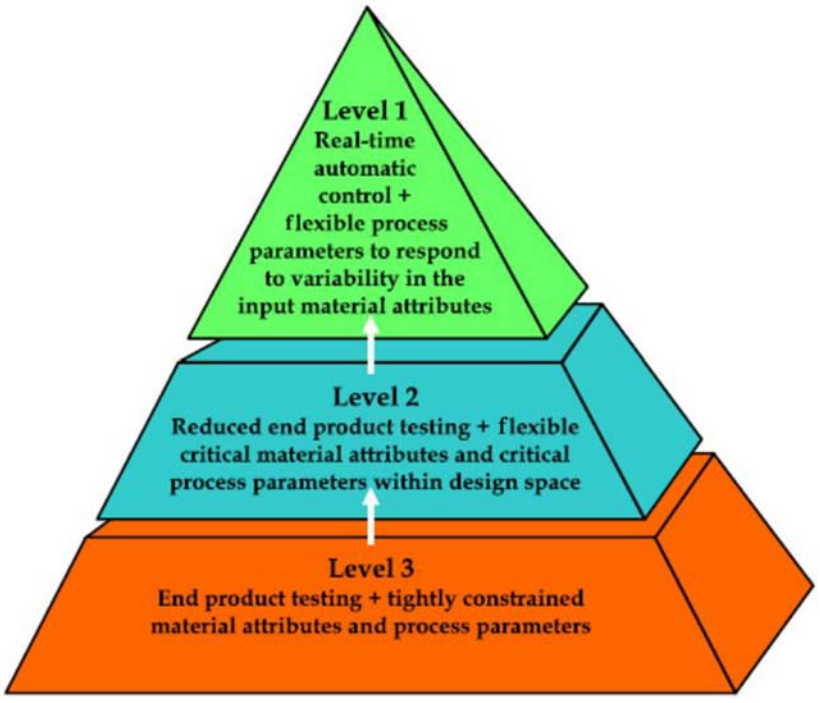


FIGURE 40.5 Control strategy implementation options.

The operating range is within the proven acceptable range, and is the routine control employed during production in order to assure reproducibility. Proven acceptable range is based on univariate experimentation and does not take into account the interaction of multiple process parameters. Proven acceptable range can provide some, but limited, information about the process.

Minimally, the control strategy should include all CMAs and CPPs; however, a more prudent approach may be to include all variables ranked as high risk in the initial risk assessment because the conclusion of the experiments is dependent on the range studied and the complex multivariate relationship between variables. In the presence of interacting CPPs and CMAs, a design space is a better approach to ensure product quality, although it is not a requirement. Design space is the multidimensional combination and interaction of input variables (eg, material attributes) and process parameters that have been demonstrated to provide assurance of quality.[4] The design space is likely established using DOE at small-scale batches and then verified at commercial scale. It is best to exploit dimensionless parameters and material attributes to define the design space because they are scale-independent. Design space is proposed by the applicant and is subject to regulatory assessment and approval. Working within the FDA-approved design space is not considered a manufacturing change. Movement out of the design space is considered to be a change, and would normally initiate a regulatory post-approval change process. Control space is within the design space, and not subject to regulatory approval. If the control space is much smaller than the design space, the process is then considered robust.

By asking about the residual risks upon implementation of the control strategy at commercial scale, the agency recognizes that in some cases, the risk of a process step to impact the drug product CQAs may only be mitigated from high risk to medium risk at the time the application is submitted. For example, an extended-release polymer-coating process can be optimized at pilot scale. However, some variability in process efficiency and polymer coating-layer thickness may be observed during scale-up from pilot to commercial scale, resulting in a slight decrease in the theoretical polymer coating level needed to achieve the target drug release profile. In this case, the impact of polymer coating process variability on drug release from tablets needs to be further monitored during validation and routine commercial manufacturing to gain additional knowledge and to facilitate continual improvement. Therefore, the risk is considered acceptable, but at a medium level in-line with current process understanding.

The additional experience and knowledge gained during routine manufacturing should be utilized for refinement of the control strategy as part of the continual improvement of the drug product during its commercial lifecycle. However, any post-approval changes in established conditions beyond the variability already provided in the application should be reported to the agency in accordance with 21 CFR 314.70 and should follow steps as outlined by guidances used for scale-up and post-approval changes.[69–74]

P.6 Reference Standards or Materials

How were the drug product reference standards obtained, certified, and/or qualified?

A reference standard is a substance prepared for use as the standard in an assay, identification, or purity test and should have a quality appropriate for its use. It is often characterized and evaluated for its intended purpose by additional procedures other than those used in routine testing.[59] In response to this question, applicants should discuss the reference standards utilized for both drug substance and degradants. If the same reference standards are used for the drug substance, it may be appropriate to reference Section 2.2.S.5. Generally, only a lot number is needed when a compendial reference standard is used as the primary reference standard. When an in-house standard is used, applicants should provide the source(s), lot number, and reference to the location of the CoA and of full characterization information in Module 3. An in-house primary reference standard should be qualified by elucidating its chemical structure using pertinent methods, such as MS, NMR, IR, UV-VIS, TGA, etc., as well as assessing its purity and should be certified by providing a certificate that gives the value of the specified property, its associated uncertainty and a statement of metrological traceability. Secondary or working standards can be qualified against primary reference standards. When USP standards are available, secondary or working standards should be qualified against the USP standards.

P.7 Container Closure System

What container closure system(s) is(are) proposed for commercial packaging of the drug product? What is the specification?

The applicant should give a description of the proposed primary container closure system for each strength of the drug product including the bottle, cap, void filler, and desiccant or the blister. A description of the label and any secondary packaging should also be provided. The applicant may give a reference to the location of the letters of authorization for Type III DMFs, as well as a reference to the location of the component specifications, technical drawings, and full testing results. The applicant should submit both in-house and

vendor CoAs with results for full testing performed to qualify the suitability of the components used to package the registration batches or a representative lot. Applicants should also submit certification per 21 CFR sections 174–186 to demonstrate that materials used in the component manufacture are safe for use in direct or indirect food contact. Note that any development studies performed to demonstrate the suitability of the selected container closure system should be submitted under the QbR question in 2.3.P.2.4.

P.8 Stability

What is the stability specification? If applicable, what is the justification for acceptance criteria that differ from the drug product release specification?

Applicants should present the stability specification in a side-by-side comparison table that highlights any differences between the release and shelf-life specification and provides justification for any differences.

What is the proposed shelf-life for the drug product? What drug product stability studies support the proposed shelf-life and storage conditions in the container closure system? How does statistical evaluation of the stability data and any observed trends support the proposed shelf-life?

Applicants may propose a shelf-life based on stability data available at the time of submission and as recommended in the relevant stability guidance documents.[75–77] The intended storage conditions to be included in the drug product labeling should be clear.

If the applicant chooses to use any bracketing or matrixing designs to support the drug product stability, then the designs should be fully explained and justified according to the considerations detailed in the applicable guidance documents.[78] Generally, 6 months of stability data at both the accelerated and long-term stability conditions are expected at the time of submission to support the proposed shelf-life. The applicant should provide a tabular summary of these data along with a reference to the location of the full stability results in Module 3. If there is a significant change observed in the accelerated data, the applicant should also submit 6 months of intermediate stability data at the time of submission. These data should be updated during the review cycle. Applicants should discuss any observed trends or out-of-specification results in the stability data and should perform appropriate statistical analysis of all quantifiable attributes in order to support any proposal for extrapolation of expiry period. Usually when significant trends or failure to comply with the acceptance criterion are observed during the accelerated study, the tentative expiration date will be determined based upon the intermediate and long-term data.

If the drug product labeling details any further manipulation or preparation prior to administration, the applicant may reference P.2.6 for studies conducted to demonstrate compatibility. In some cases, in-use stability studies should be conducted to demonstrate that the drug product quality (eg, drug product release or drug substance polymorphic form) is not impacted by patient storage in a bathroom medicine cabinet or pharmacy storage on shelf. Data to support stability in bulk packs for holding periods of more than 30 days also should be submitted. In cases where bulk drug product is transferred to another location for packaging, the applicant may consider simulated shipping studies to demonstrate that the bulk packaging is adequate to ensure the integrity of the drug product (eg, orally disintegrating tablets).

What are the post-approval stability protocols and other stability commitments for the drug product?

The applicant should submit a post-approval stability protocol for review and approval by the agency. It is a detailed plan that is used to generate and analyze stability data to support and confirm the tentative expiration date of the drug product. The post-approval stability protocol should include:

- Storage conditions
- Number of batches and packaging configurations to be entered into the stability protocol
- Testing intervals
- Tests to be performed and testing schedule

In an original submission, the applicant is also expected to make certain commitments, which constitute agreements that must be fulfilled after approval of the application. The stability commitment should include:

- The number of validation batches of the drug product, and annual production batches thereafter, to be tested for stability in accordance with the approved post-approval stability protocol. Each strength of the marketed drug product and all container closure systems should be included unless a bracketing or matrixing design was approved.
- A commitment to submit stability data in annual reports
- A description of how changes or deterioration in the distributed drug product will be handled as per 21 CFR 314.70(b)(1)(ii)

40.5 FUTURE DIRECTION

To realize the vision of OPQ, the office employs a team-based integrated quality assessment approach to maximize each team member's expertise and provide aligned patient-focused and risk-based drug quality recommendations, inclusive of drug substance, drug product, biopharmaceutics, microbiology, manufacturing, and facilities. Ideally, submissions to the agency need to be functional, meaning that they are electronic, structured, and searchable. They should also be of high quality

and demonstrate a risk- and science-based approach to drug product development. Although it seems contradictory, submissions must be both comprehensive and concise. The data should be presented in a cohesive manner that clearly tells the development story by communicating an appropriate level of detail to convey how the product and process understanding has culminated in a control strategy designed to ensure that the therapeutic benefit of the drug product promised on the label is realized by the patient. The submission should also clearly delineate what established conditions are proposed to ensure desired product quality. QbR goes a long way toward facilitating the previously mentioned needs. However, the current QbR is one-size-fits-all and was largely developed with solid oral dosage forms in mind. Therefore, future iterations may be more flexible and dosage-form specific to better focus both industry and reviewers on critical areas unique to those dosage forms.

Although this chapter focuses on the QbR questions developed for the assessment of drug product, there are also QbR questions currently used for other disciplines within OPQ, including drug substance and microbiology. The door is open for future versions of QbR to further refine the questions related to these disciplines and to add additional discipline-specific questions for biopharmaceutics and facilities. The integration of drug review, inspection, and compliance in combination with QbR tailored to each discipline will put CDER in a better position to anticipate and intervene before quality problems develop by improving the FDA's oversight of quality throughout the lifecycle of a pharmaceutical product and increasing the applicant's ability to produce high-quality products for the American public.

40.6 CONCLUSIONS

In summary, QbR questions were designed to embody QbD elements, including definition of the QTPP based on desired product performance and identification of the CQAs, systematic development studies to facilitate product and process design and gain product and process understanding, develop a comprehensive control strategy, and strive for continual improvement throughout the product lifecycle.

Quality cannot be tested into a product, but is built by design. The adoption of QbD approaches in the pharmaceutical industry is an evolving process. The goal of embracing a QbD mindset is to reduce product variability and defects, thereby enhancing product development and manufacturing efficiencies and post-approval change management. It is achieved by designing a robust formulation and manufacturing process that is understood at a mechanistic level and establishing clinically relevant specifications that are adequately controlled.

QbR establishes common submission and quality expectations and has helped to transform the CMC review into a modern, science- and risk-based pharmaceutical quality assessment. QbR is consistent with the current QbD paradigm and is congruent with risk management approaches. QbR facilitates clear communication by helping industry and the FDA speak the same "language," by clearly delineating submission expectations, and by encouraging the industry to be transparent in its thought process and to justify the choices made throughout pharmaceutical development and manufacture. Historically, industry has been reluctant to provide additional supportive information (eg, development data) within regulatory submissions to the FDA because of the presumption that providing more information will lead to more regulatory questions. From the reviewer perspective, however, the opposite is true. When the applicant does a good job responding to QbR, it conveys their product and process understanding to the agency and instills greater confidence that the drug product will meet the QTPP and deliver the promised therapeutic benefit to the patient throughout its lifecycle.

APPENDIX: QbR QUESTIONS

2.3 Introduction to the QOS

Proprietary name of drug product
Nonproprietary name of drug product
Nonproprietary name of drug substance
Proposed indication(s) including intended patient population
Duration of treatment
Maximum daily dose (MDD)
Alternative methods of administration

2.3.P Drug Product

2.3.P.1 Description and Composition

1. What is the description of the proposed commercial drug product? What are the components and composition of the final drug product as packaged and administered on both a per unit dose and %w/w basis? What is the function(s) of each excipient?
2. Does any excipient exceed the FDA IID limit for this route of administration calculated based on maximum daily dose? If so, please justify.
3. If applicable, what are the differences between this formulation and the listed/RLD formulation?

2.3.P.2 Pharmaceutical Development

4. For 505(b)(1) applications, what is the rationale for selecting the proposed dosage form for the drug product? For 505(b)(2) and 505(j) applications, what are the characteristics of the listed/RLD product? What is the QTPP of the finished product based on the proposed indication and patient populations? How is the QTPP justified?

5. What are the quality attributes of the finished product? Which quality attributes are considered CQAs? For each CQA, what is the target and how is it justified?

6. What is the approach for meeting the CQAs related to clinical performance? If applicable, what in vitro bioperformance evaluations (ie, dissolution method, flux assay, etc.) were used during pharmaceutical development to ensure clinical performance?

2.3.P.2.1 Components of the Product

2.3.P.2.1.1 Drug Substance

7. What are the physical, chemical, biological and, if applicable, mechanical properties of the drug substance, including physical description, pK_a, chirality, polymorphism, aqueous solubility (as a function of pH), hygroscopicity, melting point(s), partition coefficient and, when available, BCS classification?

8. What is the drug substance specification used to accept the incoming drug substance batches and how is it justified? For each test in the specification, provide a summary of the analytical procedure(s) and, if applicable, a summary of validation or verification report(s).

2.3.P.2.1.2 Excipients

9. What evidence supports excipient–drug substance compatibility and, if applicable, excipient–excipient compatibility?

10. What is the rationale for the excipient selections?

2.3.P.2.2 Drug Product

11. What aspects of the formulation were identified as potentially high risk to the drug product performance?

12. What formulation development studies were conducted? What attributes of the drug substance, excipients, and in-process materials were identified as critical and how do they impact the drug product CQAs?

13. How does the proposed commercial formulation differ from the formulations used during bioequivalence and/or clinical studies? What is the rationale for the formulation change? What biopharmaceutics evaluations (comparative dissolution, bioequivalence studies, biowaivers, etc.) support the formulation changes and link the development formulations to the posed commercial formulation?

2.3.P.2.3 Manufacturing Process Development

14. What is the rationale for selecting this manufacturing process for the drug product?

15. What is the potential risk of each process step to impact the drug product CQAs and how is the risk level justified?

16. For each of the potentially high-risk manufacturing unit operations:

- **a.** What input material attributes and process parameters were selected for study and what are the justifications for the selection?
- **b.** What process development studies were conducted? Provide a summary table listing batch size, process parameter ranges, equipment type, and estimated use of capacity.
- **c.** What process parameters and material attributes were identified as critical and how do they impact the drug product CQAs?
- **d.** How were the process parameters adjusted across lab, pilot/registration, and commercial scale? What are the justifications for any changes?

17. If applicable, what online/at line/in-line monitoring technologies are proposed for routine commercial production that allows for real-time process monitoring and control? Provide a summary of how each technology was developed.

2.3.P.2.4 Container Closure System

18. What specific container closure attributes are necessary to ensure drug product integrity and performance through the intended shelf life? If applicable, what are the differences in the container closure system(s) between this product and the RLD?

19. How was the container closure system(s), including bulk containers, qualified for suitability (protection, compatibility, safety, and performance)?

2.3.P.2.5 Microbiological Attributes

20. When applicable, what microbiological attributes were evaluated on the finished product?

2.3.P.2.6 Compatibility

21. If applicable, what supportive data demonstrates the compatibility of the drug product with the means of administration (eg, additives and/or diluents, other co-administered drugs, dosing device)?

2.3.P.3 Manufacture

22. Who manufactures the drug product? List each participant and facility involved in drug product manufacturing/testing activities and clearly state their function. List the date of the last FDA inspection of each facility involved and the result of the inspection. Has the manufacturer addressed all concerns raised at the FDA inspection?

23. What is the commercial batch formula and how does it differ from the registration batch formula? Provide justifications for any differences.

24. What is the flow diagram of the manufacturing process that shows all incoming materials, processing steps/unit operations, and in-process controls?

25. What is the detailed process description, including process parameters, material attributes of raw materials and intermediates, equipment type, batch size, in-process controls, including acceptance criteria and any proposed reprocessing?

26. What in-process sampling strategies and methods are used to monitor in-process material attributes that have a potential to affect quality?

27. What are the in-process test results for each process step of the registration batch(es)? What are the differences, if any, in the in-process controls for the registration batch(es) and the intended commercial batches? What are the justifications for these differences?

2.3.P.4 Control of Excipients

28. What are the excipient specifications and how are they justified? How do the proposed acceptance criteria for the material attributes of the excipients ensure the quality of the final drug product?

2.3.P.5 Control of Drug Product

29. What is the drug product specification, what is the justification, and how is it linked to the product performance and patient safety? Does the specification include all the CQAs for the drug product? Does it include all the critical drug product attributes?

30. For each test in the specification, provide a summary of the analytical procedure(s) and, if applicable, a summary of the validation or verification report(s).

31. How do the batch analysis results compare with your proposed specification? Provide a summary of the batch analysis results.

32. What are the drug product degradants? For each degradant, what is the structure, chemical name, origin, and mechanism of formation? How are the proposed limits justified and/or qualified for safety based on nonclinical studies? What is the control strategy for the potential drug product degradants?

33. What is the proposed control strategy for the drug product manufactured at commercial scale? What are the residual risks upon implementation of the control strategy at commercial scale?

2.3.P.6 Reference Standards and Materials

34. How were the drug product reference standards obtained, certified, and/or qualified?

2.3.P.7 Container Closure System

35. What container closure system(s) is(are) proposed for commercial packaging of the drug product? What is the specification?

2.3.P.8 Stability

36. What is the stability specification? If applicable, what is the justification for acceptance criteria that differ from the drug product release specification?

37. What is the proposed shelf-life for the drug product? What drug product stability studies support the proposed shelf-life and storage conditions in the container closure system? How does statistical evaluation of the stability data and any observed trends support the proposed shelf-life?

38. What are the post-approval stability protocols and other stability commitments for the drug product?

References

1. US Food and Drug Administration. Final report on pharmaceutical cGMPs for the 21st century—a risk-based approach. http://www.fda.gov/Drugs/DevelopmentApprovalProcess/Manufacturing/QuestionsandAnswersonCurrentGoodManufacturingPracticescGMPforDrugs/ucm137175.htm; Fall 2004 [accessed 19.10.15].
2. US Food and Drug Administration. Guidance for industry, Q9 quality risk management; June 2006.
3. US Food and Drug Administration. Guidance for industry, Q10 pharmaceutical quality system; April 2009.

4. US Food and Drug Administration. Guidance for industry, Q8 (R2) pharmaceutical development; November 2009.
5. US Food and Drug Administration. Center for Drug Evaluation and Research manual of policies and procedures 4000.4: clinical pharmacology and biopharmaceutics review template. Effective date: 04/27/2004.
6. US Food and Drug Administration. Center for Drug Evaluation and Research manual of policies and procedures 6025.1: good review practices. Recertified date: 02/3/2012.
7. US Food and Drug Administration. Center for Drug Evaluation and Research manual of policies and procedures 5015.10: chemistry review of question-based review (QbR) submissions. Effective date: 11/18/2014.
8. US Food and Drug Administration Office of Generic Drugs White paper on question-based review (QbR) for generic drugs: an enhanced pharmaceutical quality assessment system; 2007.
9. US Food and Drug Administration. Guidance for industry, M4Q: the CTD—quality; August 2001.
10. US Food and Drug Administration. Office of Generic Drugs. http://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/AbbreviatedNewDrugApplicationANDAGenerics/ucm120981.htm; 2009 [accessed 20.10.15].
11. Yu LX, Raw A, Lionberger R, et al. US FDA question-based review for generic drugs: a new pharmaceutical quality assessment system. *J Generic Med* 2007;**4**:239–48.
12. Yu LX. Pharmaceutical quality by design: product and process development, understanding, and control. *Pharm Res* 2008;**25**(4):781–91.
13. Yu LX, Lee L, et al. US FDA Office of Generic Drugs' pharmaceutical quality initiative: progress and feedback on question based review. *Pharm Eng* 2007;**27**:52–60.
14. Srinivasan A, Iser R. FDA perspectives: common deficiencies in abbreviated new drug applications: part 1—drug substance. *Pharm Technol* 2010;**34**(1):50–9.
15. Srinivasan A, Iser R, Gill D. FDA perspectives: common deficiencies in abbreviated new drug applications: part 2—description, composition, and excipients. *Pharm Technol* 2010;**34**(8):45–51.
16. Srinivasan A, Iser R. FDA perspectives: common deficiencies in abbreviated new drug applications: part 4—manufacture and container closure. *Pharm Technol* 2011;**35**(4):62–8.
17. Srinivasan A, Iser R, Gill D. FDA perspectives: common deficiencies in abbreviated new drug applications: part 3—control of the drug product and stability. *Pharm Technol* 2011;**35**(2):58–67.
18. Skanchy DJ. Question-based review: an FDA reviewer's perspective. *Pharm Technol* 2009;**33**(10).
19. US Food and Drug Administration. Center for Drug Evaluation and Research manual of policies and procedures 5016.1: applying ICH Q8(R2), Q9, and Q10 principles to CMC review. Effective date: 02/08/2011.
20. Inactive Ingredient Search for Approved Drug Products. Available from: http://www.accessdata.fda.gov/scripts/cder/iig/index.Cfm [accessed 05.09.15].
21. US Food and Drug Administration. Guidance for industry, ANDA submissions—refuse-to-receive standards; May 2015.
22. US Food and Drug Administration. Draft guidance for industry, applications covered by Section 505(b)(2); October 1999.
23. US Food and Drug Administration and Center for Drug Evaluation and Research, editors. Approved drug products with therapeutic equivalence evaluations; 2015.
24. US Food and Drug Administration. Guidance for industry, tablet scoring: nomenclature, labeling, and data for evaluation; March 2013.
25. US Food and Drug Administration. Guidance for industry, size, shape, and other physical attributes of generic tablets and capsules; June 2015.
26. Yu LX, Amidon G, Khan MA, Hoag SW, Polli J, Raju GK, et al. Understanding pharmaceutical quality by design. *AAPS J* 2014;**16**(4):771–83. p. Published online May 23, 2014.
27. Raw AS, Lionberger R, Yu LX. Pharmaceutical equivalence by design for generic drugs: modified-release products. *Pharm Res* 2011;**28**:1445–53.
28. US Food and Drug Administration. Quality by design for ANDAs: an example for immediate-release dosage forms. http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/AbbreviatedNewDrugApplicationANDAGenerics/UCM304305.pdf; 2012 [accessed 20.10.15].
29. US Food and Drug Administration. Quality by design for ANDAs: an example for modified-release dosage forms. http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/AbbreviatedNewDrugApplicationANDAGenerics/UCM304305.pdf; 2011 [accessed 20.10.15].
30. CMC Biotech Working Group. A mab: a case study in bioprocess development. version 2.1. www.ispe.org/pqli/a-mab-case-study-version-2.1; October 2009 [accessed 20.10.15].
31. ICH harmonised tripartite guideline: Q8(R2) pharmaceutical development; August 2009.
32. Amidon GE, He X, Hageman MJ. Physicochemical characterization and principles of oral dosage form selection. In: Abraham DJ, editor. *Burger's medicinal chemistry and drug discovery*. John Wiley & Sons, Inc. Hoboken, NJ.; 2004.
33. Amidon GL, Lennernas H, Shah VP, Crison JR. A theoretical basis for a biopharmaceutic drug classification: the correlation of in vitro drug product dissolution and in vivo bioavailability. *Pharm Res* 1995;**12**:413–20.
34. Amidon GE. Physical and mechanical property characterization of powders. In: Brittain HG, editor. *Physical characterization of pharmaceuticals solids*. New York, NY: Marcel Dekker, Inc.; 1995. p. 281–320.
35. Yu LX, Furness MS, Raw A, Woodland Outlaw KP, Nashed NE, Ramos E, et al. Scientific considerations of pharmaceutical solid polymorphism in abbreviated new drug applications. *Pharm Res* 2003;**20**:531–6.
36. Raw AS, Furness MS, Gill DS, Adams RC, Holcombe Jr. FO, Yu LX. Regulatory considerations of pharmaceutical solid polymorphism in abbreviated new drug applications (ANDAs). *Adv Drug Deliv Rev* 2004;**56**:397–414.
37. Miller SPF, Raw AS, Yu LX. FDA perspective on pharmaceutical solid polymorphism. In: Hilfiker R, editor. *Polymorphism—in the pharmaceutical and fine chemical industry*. Weinheim: Wiley-VCH, Verlag Gmbh & Co.; 2006. p. 385–403.
38. US Food and Drug Administration. Guidance for industry, ANDAs: pharmaceutical solid polymorphism chemistry, manufacturing, and controls information; July 2007.
39. US Food and Drug Administration. Guidance for industry, Waiver of in vivo bioavailability and bioequivalence studies for immediate release solid oral dosage forms based on a biopharmaceutics classification system; August 2000.
40. Amidon GE. Data driven formulation development using material sparing methods. Garnet Peck Symposium, Purdue University, September 20-21, West Lafayette, IN. 2006.
41. Rohrs BR, Amidon GE, Meury RH, Secreast PJ, King HM, Skoug CJ. Particle size limits to meet USP content uniformity criteria for tablets and capsules. *J Pharm Sci* 2006;**95**:1049–59.
42. Waterman KC, Adami RC, Alsante KM, Hong J, Landis MS, Lombardo F, et al. Stabilization of pharmaceuticals to oxidative degradation. *Pharm Dev Technol* 2002;**7**:1–32.
43. Waterman KCA, Adami RC. Accelerated aging: prediction of chemical stability of pharmaceuticals. *Int J Pharm* 2005;**293**:101–25.

44. Serajuddin ATM, Thakur AB, Ghoshal RN, Fakes MG, Ranadive SA, Morris KR, et al. Selection of solid dosage form composition through drug-excipient compatibility testing. *J Pharm Sci* 1999;**88**:696–704.
45. Mura P, Faucci MT, Manderioli A, Bramanti G, Ceccarelli L. Compatibility study between ibuproxam and pharmaceutical excipients using differential scanning calorimetry, hot-stage microscopy and scanning electron microscopy. *J Pharm Biomed Anal* 1998;**8**:151–63.
46. USP 38—NF 33 (United States Pharmacopeial Convention). *Chapter 1078. Good manufacturing practice for bulk pharmaceutical excipients*. Rockville, MD: USP; 2015. p. 1019–35.
47. USP 38—NF 33 (United States Pharmacopeial Convention). *Chapter 1059. Excipient performance*. Rockville, MD: USP; 2015. p. 977–97.
48. Nazzal S, Nutan M, Palamakula A, Shah R, Zaghloul AA, Khan MA. Optimization of a self-nanoemulsified tablet dosage form of ubiquinone using response surface methodology: effect of formulation ingredients. *Int J Pharm* 2002;**240**:103–14.
49. Awotwe-Otoo D, Agarabi C, Wu GK, Casey E, Read E, Lute S, et al. Quality by design: impact of formulation variables and their interactions on quality attributes of a lyophilized monoclonal antibody. *Int J Pharm* 2012;**438**(1–2):167–75.
50. US Food and Drug Administration. Guidance for industry, PAT—a framework for innovative pharmaceutical development, manufacturing, and quality assurance; September 2004.
51. Stahl H. Comparing different granulation techniques. http://www.gea.com/global/de/binaries/GEA_Pharma_-_Granulation-Methods_ART_GBpdf_tcm30-16923.pdf [accessed October 20.10.15].
52. Glodek M, Liebowitz S, McCarthy R, McNally G, Oksanen C, Schultz T, et al. Process robustness—a PQRI white paper. *Pharm Eng* 2006;**26**:1–11.
53. Petrides DP, Koulouris A, Lagonikos PT. The role of process simulation in pharmaceutical process development and product commercialization. *Pharm Eng* 2002;**22**:1–8.
54. Johanson JR. A rolling theory for granular solids. *ASME J Appl Mech Ser E* 1965;**32**(4):842–8.
55. Yu LX, Lionberger RA, Raw AS, D'Costa R, Wu H, Hussain AS. Application of process analytical technology to crystallization process. *Adv Drug Deliv Rev* 2004;**56**(3):349–69.
56. Zu Y, Luo Y, Ahmed SU. PAT initiative in generic product development. *Am Pharm Rev* 2007;**10**:10–16.
57. Lionberger RA, Lee SL, Lee L, Raw A, Yu LX. Quality by design: concepts for ANDAs. *AAPS J* 2008;**10**:269–76.
58. US Food and Drug Administration. Guidance for industry, container closure systems for packaging human drugs and biologics; May 1999.
59. US Food and Drug Administration. Guidance for industry, Q6A specifications: test procedures and acceptance criteria for new drug substances and new drug products: chemical substances; December 2000.
60. USP 38—NF 33 (United States Pharmacopeial Convention). *Chapter 1111. Microbiological examination of nonsterile products: acceptance criteria for pharmaceutical preparations and substances for pharmaceutical use*. Rockville, MD: USP; 2015. p. 1176–8.
61. USP 38—NF 33 (United States Pharmacopeial Convention). *Chapter 1112. Microbiological examination of nonsterile products: acceptance criteria for pharmaceutical preparations and substances for pharmaceutical use*. Rockville, MD: USP; 2015. p. 1178–80.
62. European Medicines Agency. Guideline on real time release testing (formerly guideline on parametric release); October 2012.
63. US Food and Drug Administration Center for Drug Evaluation and Research. Manual of policies and procedures 5310.7: acceptability of standards from alternative compendia (BP/EP/JP). Recertified date: 2/24/2014.
64. ICH Harmonised Tripartite Guideline, Q2(R1) validation of analytical procedures: text and methodology; November 2005.
65. USP 38—NF 33 (United States Pharmacopeial Convention). *Chapter 1226. Verification of compendial procedures.* Rockville, MD: USP; 2015. p. 1451–2.
66. US Food and Drug Administration. Guidance for industry, Q3B (R2) impurities in new drug products; July 2006.
67. ICH harmonised tripartite guideline: Q3B(R2) impurities in new drug products; June 2006.
68. US Food and Drug Administration. Guidance for industry, ANDAs: impurities in drug products; November 2010.
69. US Food and Drug Administration. Guidance for industry, changes to an approved NDA or ANDA; April 2004.
70. US Food and Drug Administration. Guidance for industry, immediate release solid oral dosage forms scale-up and post-approval changes: chemistry, manufacturing, and controls, in vitro dissolution testing, and in vivo bioequivalence documentation; November 1995.
71. US Food and Drug Administration. guidance for industry, modified release solid oral dosage forms scale-up and post-approval changes: chemistry, manufacturing, and controls, in vitro dissolution testing, and in vivo bioequivalence documentation; September 1997.
72. US Food and Drug Administration. Guidance for industry, CMC postapproval manufacturing changes to be documented in annual reports; March 2014.
73. US Food and Drug Administration. Guidance for industry, changes to an approved NDA or ANDA questions and answers; January 2001.
74. US Food and Drug Administration. Guidance for industry, SUPAC: manufacturing equipment addendum; December 2014.
75. US Food and Drug Administration. Guidance for industry, ANDAs: stability testing of drug substances and products; June 2013.
76. US Food and Drug Administration. Guidance for industry, stability testing of drug substances and products, questions and answers; May 2014.
77. US Food and Drug Administration. Guidance for industry, Q1A (R2) stability testing of new drug substances and products; November 2003.
78. ICH Harmonised Tripartite Guideline: Q1D bracketing and matrixing designs for stability testing of new drug substances and products; February 2002.

Index

Note: Page numbers followed by "*f*" and "*t*" refer to figures and tables, respectively.

D

E

G

H

I

J

K

L

M

N

O

Q

R

S

T

SECOND EDITION

Developing Solid Oral Dosage Forms

Pharmaceutical Theory & Practice

Edited by

Yihong Qiu, PhD
Senior Research Fellow, Oral Drug Products, Manufacturing Science and Technology, AbbVie Inc., North Chicago, IL, USA

Yisheng Chen, PhD
Vice President, Product Development, Novast Laboratories, Nantong, China

Geoff G. Z. Zhang, PhD
Senior Research Fellow, Drug Product Development, Research and Development, AbbVie Inc., North Chicago, IL, USA

Lawrence Yu, PhD
Deputy Director, Office of Pharmaceutical Science, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD, USA

Rao V. Mantri, PhD
Executive Director, Drug Product Science and Technology, Bristol-Myers Squibb, New Brunswick, NJ, USA

The second edition of ***Developing Solid Oral Dosage Forms: Pharmaceutical Theory and Practice*** illustrates how to develop high-quality, safe and effective pharmaceutical products by providing the latest theories, techniques, tools and scientific advances in preformulation investigation, formulation and process design, characterization and scale-up, and production operations. This book continues to cover the essential principles of physical pharmacy, biopharmaceutics and industrial pharmacy and their application to the research and development process of oral dosage forms. Chapters have been added and completely revised as necessary to produce a comprehensive, well-organized and valuable reference for industry professionals and academics engaged in all aspects of the development process. New and important topics include amorphous solid dispersion using twin-screw extrusion, spray drying, modelling and simulation, bioequivalence of complex modified-released dosage forms, biowaivers and much more.

Key Features

- Written and edited by a team of international leading experts with experience and knowledge across industry, academia and regulatory settings
- Includes new chapters covering the pharmaceutical applications of surface phenomena, predictive biopharmaceutics and pharmacokinetics, the development of formulations for drug discovery support, emerging manufacturing technologies, and much more
- Incorporates new case studies throughout the book
- Contains a section completely devoted to regulatory aspects, including global product regulation and international perspectives

Related Titles

- Di and Kerns, *Drug-Like Properties: Concepts, Structure Design and Methods from ADME to Toxicity Optimization, Second Edition,* January 2016, 9780128010761
- Kulkarni and Shaw, *Essential Chemistry for Formulators of Semi-Solid and Liquid Dosages,* June 2015, 9780128010242
- Loftsson, *Drug Stability for Pharmaceutical Scientists,* January 2014, 9780124115484

Pharmaceutical Science / Drug Development

ISBN 978-0-12-802447-8

9 780128 024478

ELSEVIER

ACADEMIC PRESS

An imprint of Elsevier

elsevier.com