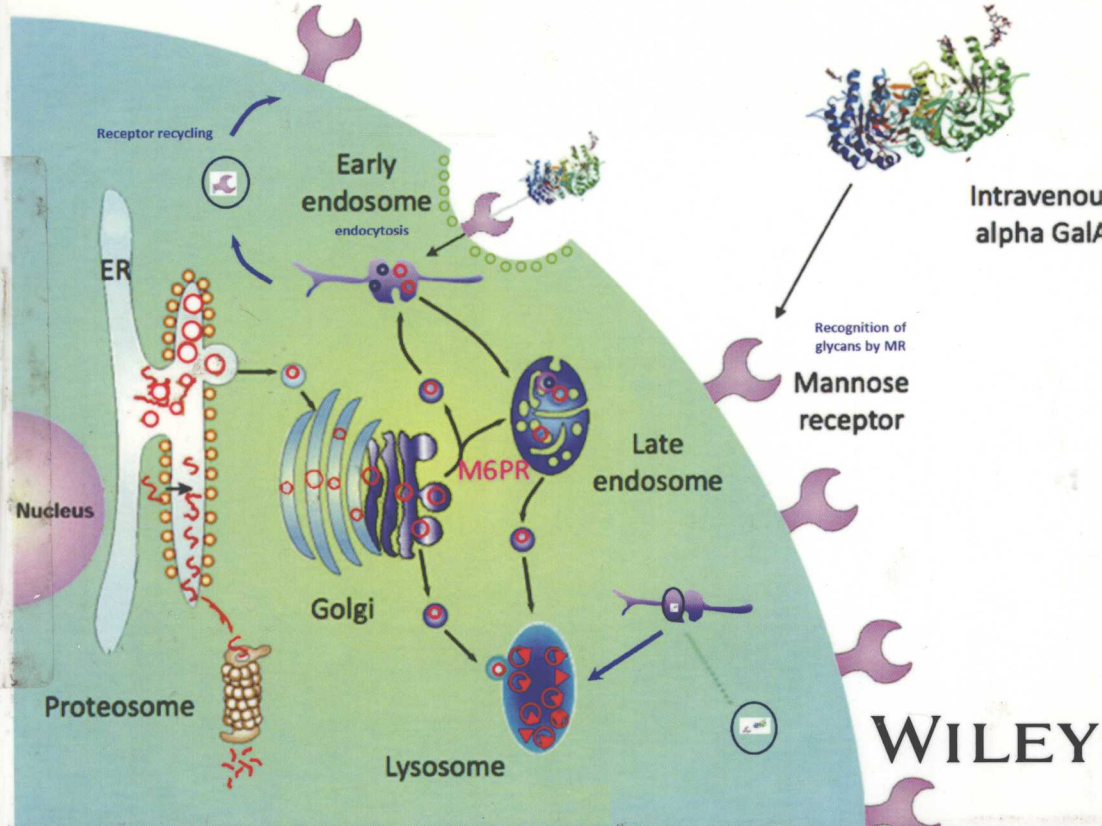
Chemical Biology of Enzymes for Biotechnology and Pharmaceutical Applications

Enzyme Technologies

Pluripotent Players in Discovering Therapeutic Agents

Edited by

Hsiu-Chiung Yang • Wu-Kuang Yeh • James R. McCarthy



WILEY

ENZYME TECHNOLOGIES

Pluripotent Players in Discovering Therapeutic Agents

Edited by

HSIU-CHIUNG YANG
Eli Lilly and Company
Indianapolis, Indiana

WU-KUANG YEH
Indiana University Purdue University Indianapolis
Indianapolis, Indiana

JAMES R. McCARTHY
Indiana University Purdue University Indianapolis
Indianapolis, Indiana

WILEY

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data:

Yang, Hsiu-Chiung.
Enzyme technologies : pluripotent players in discovering therapeutic agents / by Hsiu-Chiung Yang, Wu-Kuang Yeh, and J.R. McCarthy.
pages cm. – (Chemical biology of enzymes for biotechnology and pharmaceutical applications)
Includes index.
ISBN 978-0-470-28626-5 (hardback)
1. Enzyme inhibitors. 2. Enzymes–Pharmacokinetics 3. Drug design.
I. Yeh, Wu-Kuang, 1942– II. McCarthy, J. R. III. Title.
QP601.5.M33 2014
615.3–dc23

2013013968

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

CONTRIBUTORS

Herve Aloysius, MS, Department of Medicinal Chemistry, Ernest Mario School of Pharmacy, Rutgers, The State University of New Jersey, Piscataway, New Jersey

Michael Beck, MD, University Medical Center, Children's Hospital, Mainz, Germany

Rathnam Chaguturu, PhD, Del Shankel Structural Biology Center, High Throughput Screening Laboratory, Lawrence, Kansas; SRI International, Harrisonburg, Virginia

Yu Chen, MS, Department of Medicinal Chemistry, Ernest Mario School of Pharmacy, Rutgers, The State University of New Jersey, Piscataway, New Jersey

Ozlem Goker-Alpan, PhD, LSD Research and Treatment Unit, Center for Clinical Trials, Fairfax, Virginia

Longqin Hu, PhD, Department of Medicinal Chemistry, Ernest Mario School of Pharmacy, Rutgers, The State University of New Jersey, Piscataway, New Jersey

Daigo Inoyama, BS, Department of Medicinal Chemistry, Ernest Mario School of Pharmacy, Rutgers, The State University of New Jersey, Piscataway, New Jersey

Jutta Keller, MD, Department of Internal Medicine, Israelitic Hospital in Hamburg, Hamburg, Germany

Ivan S. Krylov, PhD, Department of Chemistry, University of Southern California, Los Angeles, California

Ley Nadine Lacbawan, PhD, LSD Research and Treatment Unit, Center for Clinical Trials, Fairfax, Virginia

Peter Layer, MD, PhD, FACG, AGAF, Department of Internal Medicine, Israelitic Hospital in Hamburg, Hamburg, Germany

Gerald H. Lushington, PhD, Molecular Graphics and Modeling Laboratory, University of Kansas, Lawrence, Kansas; LiS Consulting, Lawrence, Kansas

James R. McCarthy, PhD, Department of Chemistry and Chemical Biology, Indiana University Purdue University Indianapolis, Indianapolis, Indiana

James McGee, PhD, Quantitative Biology, Eli Lilly and Company, Indianapolis, Indiana

Charles E. McKenna, PhD, Department of Chemistry, University of Southern California, Los Angeles, California

Shujaath Mehdi, PhD, Immunoinflammation Therapeutic Strategy Unit, Sanofi Pharmaceuticals, Bridgewater, New Jersey

Taichi Ohshiro, PhD, Department of Microbial Chemistry, Graduate School of Pharmaceutical Sciences, Kitasato University, Tokyo, Japan; Section on Lipid Sciences, Department of Pathology, Wake Forest University School of Medicine, Winston-Salem, North Carolina

Richard G. Peterson, PhD, PreClinOmics, Inc., Indianapolis, Indiana

Anuradha Roy, PhD, Del Shankel Structural Biology Center, High Throughput Screening Laboratory, Lawrence, Kansas

Henrike von Schassen, MD, Department of Internal Medicine, Israelitic Hospital in Hamburg, Hamburg, Germany

Hiroshi Tomoda, PhD, Department of Microbial Chemistry, Graduate School of Pharmaceutical Sciences, Kitasato University, Tokyo, Japan

Yanhui Yang, PhD, Department of Medicinal Chemistry, Ernest Mario School of Pharmacy, Rutgers, The State University of New Jersey, Piscataway, New Jersey

Wu-Kuang Yeh, PhD, Indiana University Purdue University Indianapolis, Indianapolis, Indiana

Wei Zheng, PhD, National Center for Advancing Translational Sciences, National Institutes of Health, Bethesda, Maryland

PREFACE

The human genome was predicted to contain approximately 2742 genes that encode enzymes, which corresponds to 9.5% of the genome (HumanCyc version 7.5). These predicted enzymes can be subdivided into 1653 metabolic enzymes and 1089 nonmetabolic enzymes (including enzymes whose substrates are macromolecules, such as protein kinases and DNA polymerases). Enzymes play an important role in human physiology and the pathophysiology of disease. Understanding the function of an enzyme presents a significant opportunity for finding therapeutic agents. Therefore, a comprehensive understanding of all aspects of enzyme technology is critical in discovering therapeutic agents targeting enzymes, therapeutic enzymes, and enzyme-based applications, enabling research in drug discovery, chemistry, material science, and a vast number of other fields in science and technology.

This series on Chemical Biology of Enzymes for Biotechnology and Pharmaceutical Applications consists of three volumes. Volume I, *Enzyme Technologies: Metagenomics, Evolution, Biocatalysis, and Biosynthesis*, was published in 2010. Volume II, *Enzyme Technologies in Drug Discovery*, as listed in Volume I, has now been changed to *Enzyme Technologies: Pluripotent Players in Discovering Therapeutic Agents*. This book is intended both for biotech and pharmaceutical scientists in academic research institutes and industry as a comprehensive reference material for all common applications of enzyme technology in drug discovery. This book will be useful for biotechnology, biochemistry, molecular biology, and medicinal chemistry faculty members for teaching or conducting research in the field of enzyme technology. It is also a practical handbook for industrial scientists to study various aspects of enzyme technology and discover new treatments for unmet medical needs.

This book is divided into three parts: Part A: Enzymes – Essential Workhorses in Pharmaceutical Research; Part B: Enzymes – Indispensable Tools for Improving Druggability; and Part C: Enzymes – Powerful Weapons for Correcting Nature's Errors. Part A consists of four chapters. Chapter 1, by Dr. A. Roy et al., discusses the principles of assay development and cutting-edge technologies available for protease assays, using proteases as a prototype. Chapter 2, by Drs. Ohshiro and Tomoda, provides a case study on the design and development of selective enzyme inhibitors, using lipid metabolizing enzymes as a prototype. There is a belief that covalent enzyme inhibitors (also called "irreversible inhibitors") are not desirable for drug candidates. Chapter 3, by Dr. Mehdi, describes methods for characterizing covalent inhibitors and their therapeutic applications and explains how enzyme kinetics has been applied in drug discovery. Chapter 4, by Drs. Yeh and Peterson, provides a comprehensive coverage on various technologies that have been applied for *in vitro* enzymatic assays, as well as on common *in vivo* models to assess preclinical drug discovery for metabolic diseases. After going through this part, readers will have a better understanding as to how to select the best enzyme targets for drug discovery, the steps involved in designing enzyme inhibitors for therapeutic agents, and methods for evaluating selective enzyme inhibitors.

Part B consists of three chapters. It explains the principles of improving druggability and provides examples on how to utilize the properties of enzymes for designing therapeutic agents, specifically prodrugs. Chapter 5, by Dr. Hu et al., provides a comprehensive review of enzymes that are being, or can be, used to design prodrugs to improve druggability of existing drug molecules. Chapters 6 and 7 summarize case studies on the design of two successful prodrugs. Chapter 6, by Dr. McCarthy, provides a detailed approach and explains the hypothesis and rationale for the design and synthesis of a Gemcitabine prodrug. Chapter 7, by Drs. McKenna and Krylov, presents several examples of successful prodrug approaches that explicitly depend on enzyme-mediated activation. After going through this part, readers will have a complete understanding on how to select best target molecules for the prodrug approach to improve druggability and how to design successful prodrugs.

Part C consists of three chapters and provides a different viewpoint as to how enzymes can be used in pharmaceutical applications. There are many types of genetic in-born error disorders. Some of these disorders are involved in either deficiency or malfunctioning of specific enzymes. In such cases, functional enzymes can be reintroduced into patients. With recent advances in protein technology, there are several successful examples, such as the production of high-quality proteins and optimal methods for the delivery of these large molecules to the body. However, further improvements on either manufacturing or delivery of enzymes for therapeutics are still required. Chapter 8, by Dr. Beck, deals with Hunter's syndrome, while Chapter 9, by Dr. Lacbawan et al., discusses enzyme replacement therapy for Fabry disease. Chapter 10, by Dr. von Schassen et al., analyzes the level and activity of pancreatic enzymes as a means of diagnosing patients with pancreatic dysfunction, for example, Hunter's syndrome. After

going through this part, readers will have an insight into how enzymes can be applied as therapeutic agents or diagnostic tools. Furthermore, readers may be able to identify possible enzyme targets to treat genetic disorders that still do not have effective medication.

Eli Lilly and Company
Indiana University Purdue University Indianapolis
Indianapolis, Indiana

HSIU-CHIUNG YANG
WU-KUANG YEH
JAMES R. MCCARTHY

CONTENTS

PART A

ENZYMES – ESSENTIAL WORKHORSES IN PHARMACEUTICAL RESEARCH

1

ASSAY TECHNOLOGIES FOR PROTEASES

ANURADHA ROY

Del Shankel Structural Biology Center, High Throughput Screening Laboratory, Lawrence, Kansas

GERALD H. LUSHINGTON

Molecular Graphics and Modeling Laboratory, University of Kansas, Lawrence, Kansas; LiS Consulting, Lawrence, Kansas

JAMES MCGEE

Quantitative Biology, Eli Lilly and Company, Indianapolis, Indiana

RATHNAM CHAGUTURU

Del Shankel Structural Biology Center, High Throughput Screening Laboratory, Lawrence, Kansas; SRI International, Harrisonburg, Virginia

I. INTRODUCTION

Proteases are ubiquitously expressed enzymes which catalyze hydrolysis of peptide bonds and work under a wide range of conditions using diverse catalytic mechanisms [1]. Proteases specifically cleave protein substrates either from the N- or C-termini (aminopeptidases and carboxypeptidases, respectively) or in the middle of the molecule (endopeptidases) [2]. Proteolytic enzymes modulate many physiological processes ranging from nonspecific hydrolysis of dietary proteins to highly specific and regulated proteolysis in cell cycle regulation, tissue remodeling,

Enzyme Technologies: Pluripotent Players in Discovering Therapeutic Agents, First Edition.
Edited by Hsiu-Chiung Yang, Wu-Kuang Yeh, and James R. McCarthy.

blood coagulation, blood pressure control, angiogenesis, apoptosis, inflammation, ovulation, fertilization, and embryonic development [3,4]. Over 500 proteases each from humans, rat, mouse, and chimpanzee have been annotated and compiled in the Degradome database (http://degradome.uniovi.es) [5,6]. Information on all known proteases and their substrates/inhibitors is listed in the MEROPS database [7]. Based on the amino acid or metal that catalyzes the nucleophilic attack on substrate peptide bonds, the proteases are classified into five major types: aspartic (Asp), metallo-, cysteine (Cys), serine (Ser), and threonine (Thr) proteases. Aspartic and metalloproteases use an activated water molecule as a nucleophile to attack the peptide bond of the substrate, whereas in Cys, Ser, and Thr proteases, a catalytic amino acid residue (Cys, Ser, or Thr, respectively) serves as a nucleophile (Fig. 1). As a result, acyl–enzyme intermediates are formed only in the reactions catalyzed by Ser/Thr and Cys peptidases. Within each class of protease type are several enzymes that may have overlapping or distinct substrate recognition sites. Rawlings and Barrett proposed a classification of proteases into families based on amino acid sequence similarity, and families with similar three-dimensional folding are assembled into clans, indicating common ancestry [7,8]. The focus of this article is mainly on mammalian proteases and retroviral proteases which are of significant therapeutic relevance.

While pepsin in gastric juices digests a variety of proteins with broad specificity, renin is an example of Asp protease that shows high substrate specificity. Most proteases bind their substrates in fairly similar manner, first elucidated for papain by Schechter and Berger [9–11]. The catalytic site is flanked on one or both sides by sites that confer specificity of substrate binding to the protease and accommodate a side chain of an amino acid residue of the substrate. The enzymatic binding sites toward the N-terminus of the substrate are the non-prime side designated as S1, S2, ..., Sn from the catalytic site, and the residues C-terminal to the cleavage site are the prime side designated as S1′, S2′, ..., Sn′ [8,12,13]. The amino acid residues in the protein substrate which correspond to their respective subsites are numbered P1, P2, ..., Pn and P1′, P2′, ..., Pn′ (Fig. 1). Only few of the substrate binding sites have stringent specificities. For instance, site S1 confers specificity for Ser proteases and caspases, whereas the site S2, a hydrophobic subsite, defines specificity for the papain family of Cys proteases. In addition to the sites close to the catalytic site of the enzyme, distant sites on the enzyme may also contribute to the binding of substrates to the protease [9]. The specificity and biological activity of caspases are also determined by S4, which is distant from the catalytic site [14]. Proteolytic processing is being recognized as a mechanism for regulation of enzymatic activities, localization, and fate of proteins that are activated by limited and specific hydrolysis of peptide bonds. Dysregulation of proteolytic activity, structure, or expression results in major pathologies in the areas of cardiovascular diseases, cancer, neurodegenerative disorders, osteoporosis, diabetes type II, pancreatitis, inflammation, arthritis, and infectious diseases [4]. A large number of marketed drugs target the proteolytic enzymes that are involved in pathogenesis of various diseases [15] (Table 1). Although only a relatively small number of proteases are currently targeted for

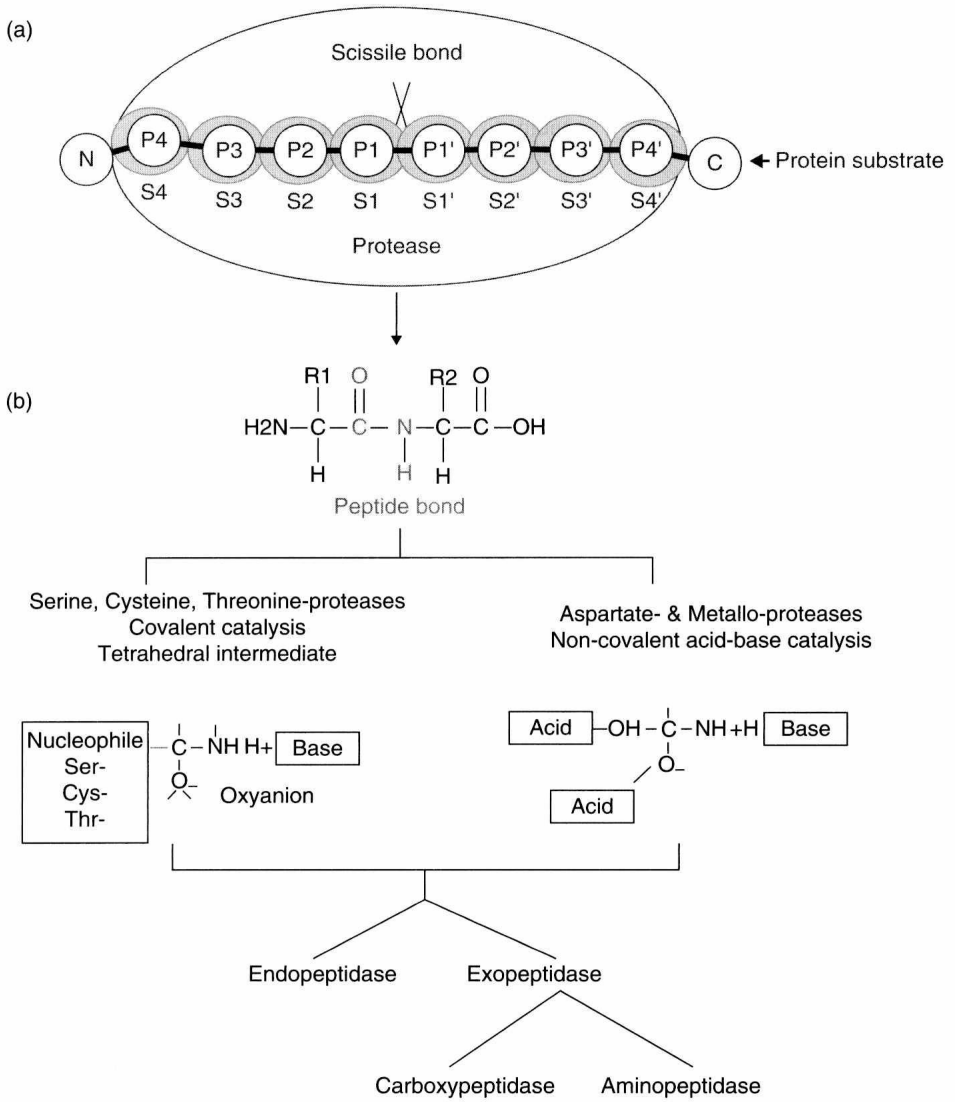


FIGURE 1 Schematic representation of binding of substrate to a protease site. (a) The binding sites of the protease are numbered on either side of the scissile bond, with the non-primed sites (S1, S2, …, S*n*) located toward the amino-terminus of the substrate and S1′ … S*n*′ or the primed subsites toward the carboxy-terminus. (b) Structure of the peptide bond which is hydrolyzed by proteases and the two basic catalytic mechanisms for all types of protease hydrolysis. In Ser, Cys, and Thr proteases, an amino acid at the active site serves as the nucleophile forming a transient covalent intermediate, whereas in metallo- and Asp proteases, an active water molecule functions as nucleophile (adapted from Reference [13]). The base in covalent catalysis is usually a His, and in non-covalent intermediate, Asp/Glu and zinc (metalloproteinases) serve as acids and bases. The proteases are also classified as endo- and/or exo-proteinases based on their ability to cleave within or at the amino-/carboxy-terminus of the peptide chain.

TABLE 1 FDA-Approved Drugs for Select Proteases

Protease	Class	Compound	Company	Indication
HIV-1 protease	Asp	Atazanavir	Bristol-Myers Squibb	AIDS
		Darunavir	Prezista	
		Fosamprenavir	GlaxoSmithKline	
		Indinavir	Merck	
		Lopinavir	Abbott	
		Nelfinavir mesylate	Pfizer	
		Ritonavir	Abbott	
		Saquinavir	Hoffmann-La Roche	
		Tipranavir	Boehringer Mannheim	
Renin	Asp	Aliskiren (Tekturna)	Novartis	Hypertension
ACE	Metallo	Captopril	Bristol-Myers Squibb	Hypertension
		Enalapril	Merck	
		Lisinopril	AstraZeneca	
Carboxypeptidase A	Metallo	Penicillamine		Hypertension
MMP-1 and MMP-2 collagenases	Metallo	Periostat	Galderma Labs	Periodontitis
Enkephalinase	Metallo	Racecadotril		
Thrombin	Ser	Ximelagatran	AstraZeneca	Thrombosis
		Argatroban	Mitsubishi Pharma	
		Lepirudin	Aventis	
		Desirudin	Novartis	
Human FXa	Ser	Fondaparinux	Sanofi-Synthélabo	Thrombosis
Human neutrophil elastase	Ser	Sivelestat	Ono	Respiratory disease
Trypsin-like protease	Ser	Camostat mesilate	Ono	Pancreatitis I
Broad-spectrum protease	Ser	Nafamostat mesilate	Japan Tobacco	Pancreatitis inflammation
Plasminogen activator	Ser	Streptokinase		
Proteasome	Thr	Bortezomib (Velcade)	Millennium	Multiple myeloma
DPPIV	Ser	Pioglitazone	Takeda	Diabetes mellitus type II
		Saxagliptin	Bristol-Myers Squibb	
		Linagliptin	Boehringer Ingelheim	
Cathepsin K	Cys	Odanacatib	Merck	Osteoporosis/ bone cancer
		ONO5334	Ono	Osteoporosis

drug development, the commercial success of angiotensin-converting enzyme (ACE) inhibitors and human immunodeficiency virus (HIV) protease inhibitors makes the protease family a valuable target for disease treatment [16]. We have worked on a variety of protease assay formats using proprietary substrates for HTS and will present an overview of common protease assay technologies. We will end the review by discussing computational approaches to designing substrates for protease binding sites.

II. PROTEASE ACTIVITY ASSAYS

Identification of appropriate substrate is the first major step toward characterizing a protease and developing an assay to monitor its activity. A protein containing a short recognition sequence for a protease may work well in an *in vitro* assay, but may not be a physiologically relevant substrate for the proteolytic enzyme. In cases where the peptide sequences are unknown or ill defined, positional scanning experiments or phage display methods are used for screening combinatorial libraries of peptides for specificity determination. Bioinformatics-based prediction of cleavage sites and determination of protease preferences on synthetic substrates are important techniques in predicting natural protease substrates. Figure 2 summarizes the techniques employed in mapping substrates, a rapidly evolving field which is beyond the scope of this article and has been extensively reviewed [17–20]. The design of sensitive and selective synthetic peptide-substrate cleavage assays and the comprehensive mapping of active site specificity determinants are crucial for developing protease inhibitor drugs. The assays for proteolytic enzymes are generally continuous, homogeneous assays that can be performed in medium- to high-throughput formats. The assays are more relevant when the enzyme activity is measured with their specific native substrates, but assays with long native proteins have low hydrolytic rates and are time-consuming and costly. The substrate peptide sequences for most of the common proteases are very well established, and the fact that most proteases catalyze hydrolysis of small peptides has led to the development of technically simple and sensitive assays using fluorometric, colorimetric, and bioluminescent methods in which a single specific peptide bond is cleaved and the cleavage is monitored spectrometrically. A large number of kits are available from various vendors that serve to assay protease families using generic substrate peptides. In general, the minimal defined peptide substrate (average three to six amino acid residues) is synthesized based on the information on the binding fragments of natural substrates or inhibitors of proteases. In the case of many proteases like caspases, matrix metalloproteinases (MMPs), cathepsins, or HIV-1 proteases, the same substrate peptide or a minor variant of a substrate sequence is conjugated with either fluorophores or chromogenic groups or tagged with aminoluciferin to allow development of fluorescence-, absorbance-, or luminescence-based assay formats. Both cell-based and rapid mix-and-read biochemical methods have been developed for assaying the protease activity. The biochemical assays based on purified recombinant enzymes being