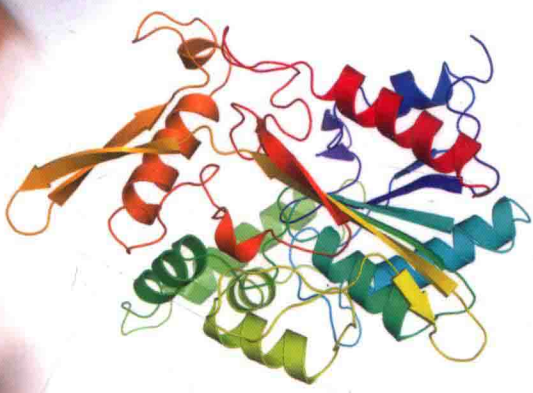
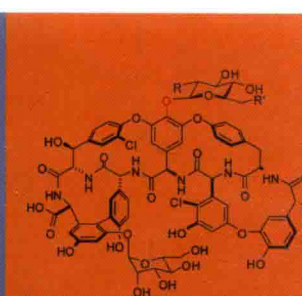


Editors
Junhua Tao
Guo-Qiang Lin
Andreas Liese

Biocatalysis for the Pharmaceutical Industry

Discovery, Development, and Manufacturing

WILEY



BIOCATALYSIS FOR THE PHARMACEUTICAL INDUSTRY
Discovery, Development, and Manufacturing

Junhua (Alex) Tao

Elevance Renewable Sciences, USA

Guo-Qiang Lin

Shanghai Institute of Organic Chemistry, CAS, China

Andreas Liese

Hamburg University of Technology, Germany

John Wiley & Sons (Asia) Pte Ltd

Other Wiley Editorial Offices

John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA

Wiley-VCH Verlag GmbH, Boschstrasse 12, D-69469 Weinheim, Germany

John Wiley & Sons Australia Ltd, 42 McDougall Street, Milton, Queensland 4064, Australia

John Wiley & Sons Canada Ltd, 5353 Dundas Street West, Suite 400, Toronto, ONT, M9B 6H8, Canada

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Library of Congress Cataloging-in-Publication Data

Tao, Junhua.
Biocatalysis for the pharmaceutical industry: discovery, development, and manufacturing/Junhua Tao, Guo-Qiang Lin, Andreas Liese.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-470-82314-9 (cloth)
1. Pharmaceutical biotechnology. 2. Enzymes–Biotechnology. I. Lin, Guo-Qiang. II. Liese, A. (Andreas) III. Title.
RS380.T36 2009
615′.19–dc22 2008045613

ISBN 978-0-470-82314-9 (HB)

Typeset in 10/12pt Times by Thomson Digital, Noida, India.
Printed and bound in Singapore by Markono Print Media Pte Ltd, Singapore.
This book is printed on acid-free paper responsibly manufactured from sustainable forestry in which at least two trees are planted for each one used for paper production.

BIOCATALYSIS FOR THE PHARMACEUTICAL INDUSTRY

Preface

Biocatalysis is evolving to be a transformational technology as a result of a confluence of factors, which include (1) large scale and ever increasingly cost-efficient DNA sequencing technologies; (2) exponential growth in GenBank; (3) powerful directed enzyme evolution and high-throughput screening technologies; (4) robust expression systems for enzyme production; (5) deep understanding of the logic of natural product biosynthesis; (6) industrial successes of metabolic engineering and pathway engineering.

Consequently, many successful stories and a number of reviews have been reported recently in developing biocatalysis for the pharmaceutical industry, across drug discovery, development, and manufacturing. The book is dedicated to these advances, and divided into four parts:

- Chapters 1–4 serve as an introduction to emerging biocatalysts, modern expression hosts, state of the art of directed evolution, high-throughput screening, and bioprocess engineering for industrial applications.
- Chapters 5–8 are directed to emerging enzymes, which include oxynitrilases, aldolases, ketoreductases, oxidases, nitrile hydratases, and nitrilases, and their recent applications especially in synthesis of chiral drugs and intermediates.
- Chapters 9 and 10 focus on synthesis of drug metabolites and intermediates catalyzed by P450s or whole cells.
- Chapters 11–13 are devoted to combinatorial biosynthesis, metabolic engineering, and autonomous enzymes for the synthesis and development of complex medicinal molecules.
- Chapter 14 discusses the recent impact of biocatalysis in green chemistry and chemical development.

Our main goal is to come up with a concise but comprehensive, practical but insightful book covering the topics discussed above. We hope you enjoy reading this book. Any suggestions and comments are welcome.

Junhua (Alex) Tao
Elevance Renewable Sciences, USA
Email: Junhua_tao@yahoo.com

Guo-Qiang Lin
Guo-Qiang Lin Shanghai Institute of Organic Chemistry, China
Email: lingq@mail.sioc.ac.cn

Andreas Liese
Hamburg University of Technology, Germany
Email: liese@tuhh.de

Contents

1

Enzymes and Their Synthetic Applications: An Overview

Junhua (Alex) Tao[1] and Jian-He Xu[2]

[1]*Elevance Renewable Sciences, 175 E. Crossroads Parkway, Bolingbrook, IL 60440, USA*
[2]*East China University of Science and Technology (ECUST) 130 Meilong Road, Shanghai, 200237, PR China*

1.1 Introduction

Whole-cell biocatalysis has been exploited for thousands of years; for example, in preparing barley for beer brewing. While the chemical, economic and social advantages of biocatalysis over traditional chemical approaches were recognized a long time ago, their applications for the drug industry have been largely underexplored until the recent technological breakthroughs in large-scale DNA sequencing, robust protein expression systems, metabolic engineering and directed evolution. In this chapter, emphasis will be directed to the discussion of those isolated enzymes which are uniquely suited for the synthesis of small-molecule pharmaceutical ingredients.

1.2 Enzyme Families

Based on reactions they catalyze, enzymes can be broadly classified into six major categories (Table 1.1) [1]. It was estimated that about 60% of biotransformations currently rely on the use of hydrolases, followed by 20% of oxidoreductases [2]. On the other hand, some of the C–C bond-forming and oxygenation enzymes catalyze reactions with very high reaction efficiency and very low waste generation, underlining the potential of emerging enzymes.

Biocatalysis for the Pharmaceutical Industry edited by Junhua Tao, Guo-Qiang Lin, and Andreas Liese

Table 1.1 Enzyme classes

Enzyme class	Examples	Reaction catalyzed
Hydrolases	lipase, protease, esterase nitrilase, nitrile hydratase glycosidase, phosphatase	hydrolysis reactions in H_2O
Oxidoreductases	dehydrogenase, oxidase oxygenase, peroxidase	oxidation or reduction
Transferases	transaminase, glycosyltransferase transaldolase	transfer of a group from one molecule to another
Lyases	decarboxylase, dehydratase, deoxyribose-phosphate aldolase	nonhydrolytic bond cleavage
Isomerases	racemase, mutase	intramolecular rearrangement
Ligases	DNA ligase	bond formation requiring triphosphate

1.3 Enzyme Discovery and Optimization

Traditionally, enzymes are discovered through screening of environmental samples and culture enrichment. As a result of recent technological breakthroughs in large-scale DNA sequencing and high-throughput screening, both the metagenomic approach and sequence-based discovery have drastically shortened the cycle of enzyme discovery.

In the metagenomic approach, DNA was directly extracted from uncultured samples followed by cloning and expression [3]. For example, by combination of directed evolution with the metagenome approach, an α-amylase mutant with optimal activity at pH 4.5 and optimal thermostability at 105 °C was discovered for starch liquefaction and EtOH production [4].

Sequence-based discovery (genome hunting) is increasingly attractive, as the public sequence bank is growing rapidly. In this approach, known sequences encoding an enzyme of interest are used to search gene databases to uncover enzymes of homologous sequences. For example, using this method, a library of deoxyribose-phosphate aldolases (DERAs) were rapidly constructed and from them a novel DERA was identified to catalyze a sequential aldol reaction of a nonnative substrate with high throughput and excellent stereoselectivity for the synthesis of statin side chains [5].

Since most synthetic applications require enzymes catalyzing nonnatural substrates, their properties often have to be improved. One way to achieve this is to optimize reaction conditions such as pH, temperature, solvents, additives, etc. [6–9]. Another way is to modulate the substrates without compromising the synthetic efficiency of the overall reaction [10]. In most cases for commercial manufacturing, however, the protein sequences have to be altered to enhance reactivity, stereoselectivity and stability. It was estimated that over 30 commercial enzymes worldwide have been engineered for industrial applications [11]. Precise prediction of which amino acids to mutate is difficult to achieve. Since the mid 1990s, directed evolution

has been demonstrated to be a powerful and robust technology to improve the desired properties [12,13]. Among them, the error-prone PCR method is probably the most popular to create random mutants by changing polymerization reaction conditions [14]. Alternatively, recombination of homologous sequences or DNA-shuffling methods can be used to introduce mutants with improved properties [15]. The major challenge in directed evolution is not generation of mutant libraries; rather, it is the availability of high-throughput assays [16]. In most cases, it requires screening of tens of thousands of mutants, which is usually tedious and time consuming. As more and more protein structures are available from the protein database, focused directed evolution or semi-rational protein design is becoming more and more popular [17]. In this approach, the three-dimensional (3-D) structure of a suboptimal enzyme is constructed by a computer algorithm from a homologous enzyme with known 3-D structure. Docking studies are then applied to search potential 'hot spots', which are then swapped with other amino acids by site-saturation mutagenesis. In this way, there are generally less than a few thousand mutants to be screened, significantly shortening the cycle of enzyme development. The fact that most beneficial mutations are proved to be near the active site makes this approach even more attractive [18].

1.4 Enzyme Production

Although some enzymes are still extracted from animal or plant tissue, most of them are now produced from microorganisms by fermentation. Bacteria and fungi are the most popular hosts for producing industrial enzymes, due to easy handling and high productivity. They can also be readily genetically engineered to improve their performance; for example, by incorporating secretion systems to facilitate enzyme isolation and purification. Some of the most popular expression hosts are *Escherichia coli*, *Pichia pastoris*, *Pseudomonas fluorescens*, *Aspergillus* sp. and *Bacillus* sp. Mammalian or plant cells are used in special cases [19–21]. By regulation, the production host should have GRAS status (Generally Regarded as Safe Status).

In a typical enzyme production procedure, cells containing genes encoding desired enzymes are grown in an Erlenmeyer flask. At large scale, a computer-controlled fermenter or bioreactor is required to maintain an appropriate control of pH, O_2, NH_3 and CO_2 to maximize cell density. The cells are harvested by centrifugation in a batch or continuous fashion. Alternatively, they can be collected through membrane filtration devices. The cell membranes are then disrupted by an ultrasonicator or French press at small scale. At a scale of over 5–10 L, a homogenizer is usually used. After centrifugation to remove cell debris, the crude enzymes remain in the supernatant and can be concentrated through precipitation by adding either inorganic salts (e.g. ammonium sulfate) or organic solvents (e.g. acetone). The crude enzymes are then purified by dialysis or a variety of chromatographic methods. The dry powder is usually obtained after lyophilization under freeze-drying conditions [22,23].

1.5 Enzymes and Synthetic Applications

Historically, the most popular enzymes used for chemical synthesis are lipases, esterases, proteases, acylases and amidases, among others. Recently, a number of recombinant biocatalysts have been discovered and isolated, significantly expanding the toolbox for biotransformations. In this section, the focus will be on these new enzymes.

1.5.1 Ketoreductases (EC 1.1.1.2)

Ketoreductases (KREDs) catalyze the conversion of a wide range of ketones and some aldehydes to chiral alcohols regio- and stereo-selectively in the presence of NADH or NADPH (Figure 1.1) [24,25]. This powerful transformation has been demonstrated in a number of industrial transformations using either isolated enzymes or whole cells. The use of isolated enzymes is often preferred because of a higher volumetric productivity and the absence of side reactions. A key to its success is the availability of efficient and cost-effective cofactor regeneration methods by using a formate dehydrogenase to recycle NAD^+ or a glucose dehydrogenase to recycle $NADP^+$ (Figure 1.1) [26,27]. It shall be noted that some alcohol dehydrogenases are also able to catalyze the oxidation of alcohols to ketones or aldehydes [28].

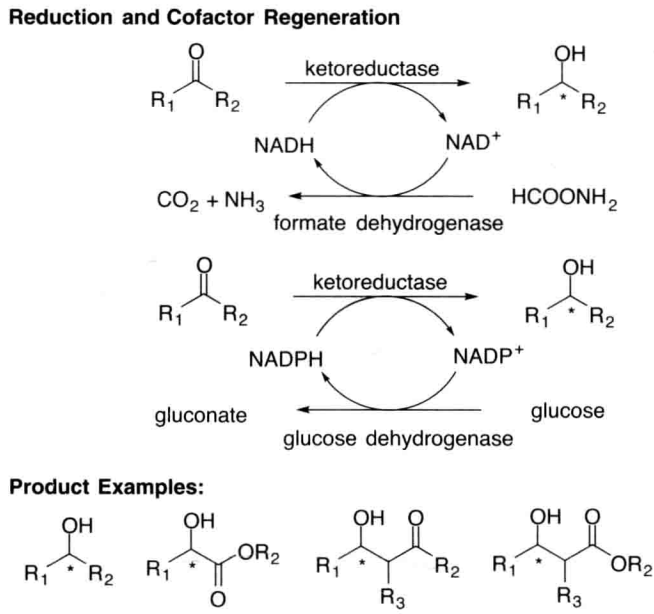


Figure 1.1

1.5.2 Enoate Reductases or Ene Reductases (EC 1.3.1.16)

Enoate reductase (ER) catalyzes NAD(P)H-dependent reduction of carbon–carbon double bonds of nonactivated enoates, as well as of α,β-unsaturated aldehydes, ketones, nitros, and nitriles (Figure 1.2). For example, the ER from *Clostridium tyrobutyricum* shows high stereospecificity and regioselectivity and broad substrate specificity [29]. Alkanes with up to two chiral centers can be directly produced by asymmetric reduction of electron-deficient alkene enzymes from the 'old yellow enzyme' family at the expense of NAD(P)H. The cofactor can be regenerated *in vitro* using a formate dehydrogenase or glucose dehydrogenase. Alternatively, a whole-cell system can be used to co-express ERs with redox enzymes for NAD(P)H recycling [30].

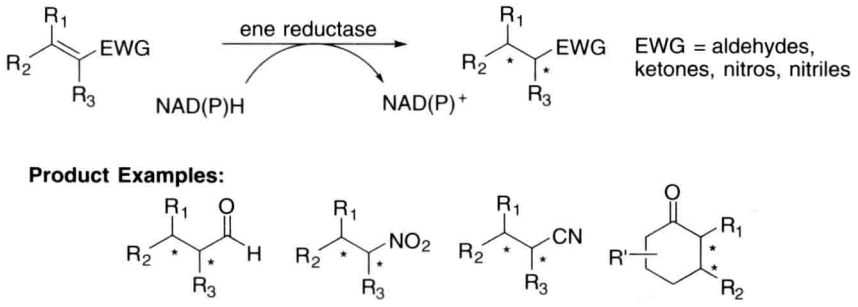


Figure 1.2

1.5.3 Oxygenases (EC. xxxx)

Oxygenases catalyze direct incorporation of molecular oxygen into substrates to produce oxygenated molecules [31,32]. They are categorized as either monooxygenases (MOs) or dioxygenases, depending upon whether one or both atoms of dioxygen are inserted into a substrate. The metal-dependent MOs, such as P450s, catalyze a wide range of reactions via metal oxo species (e.g. hydroxylation of alkanes and aromatics; epoxidation of alkenes (Figure 1.3)), while flavin adenine dinucleotide (FAD)-dependent MOs are found to catalyze oxidation of heteroatoms (S, N, Se, P) and Baeyer–Villiger reactions via FAD-hydroperoxide (FAD-OOH) species (Figure 1.4).

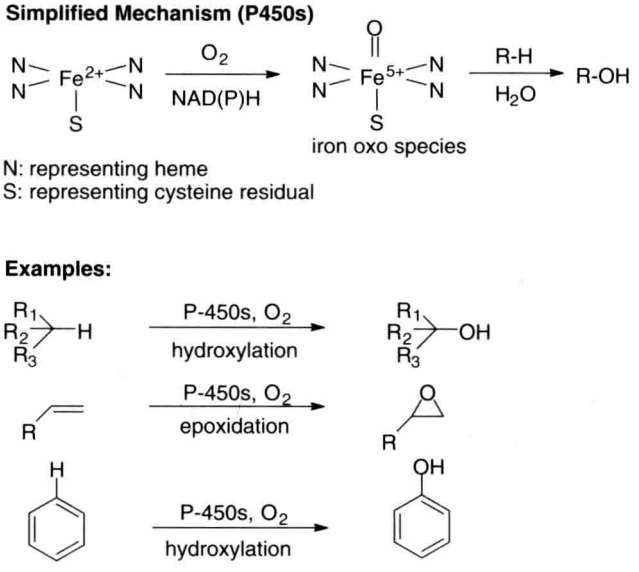


Figure 1.3

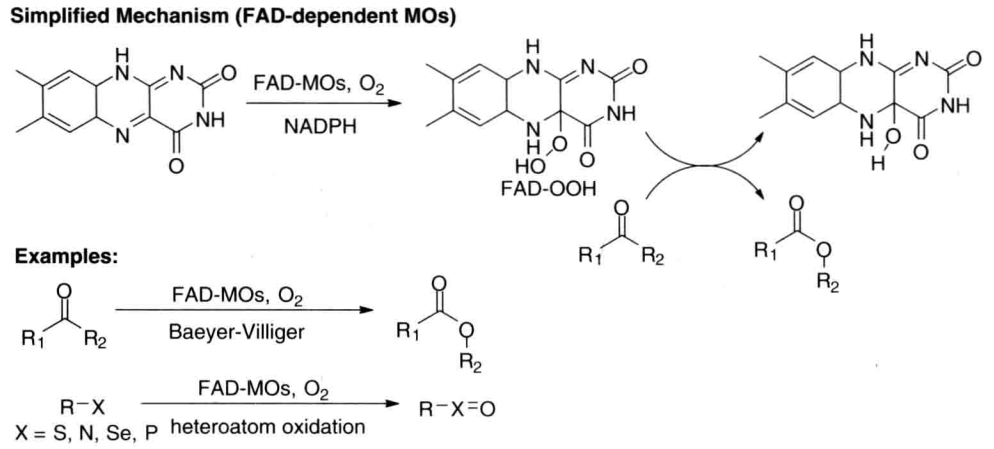


Figure 1.4

Dioxygenases usually contain a tightly bound iron atom and catalyze hydroperoxidation of allylic molecules or carboxylic acids, and dihydroxylation of aromatics (Figure 1.5) [33].

Currently, these oxygenation reactions are usually carried out in whole cells, the outcome of which is often unpredictable. The discovery of novel oxygenases and efficient hosts for protein expression remain keys to further expanding the applications of these enzymes in chemical synthesis and drug metabolism studies [34–37].

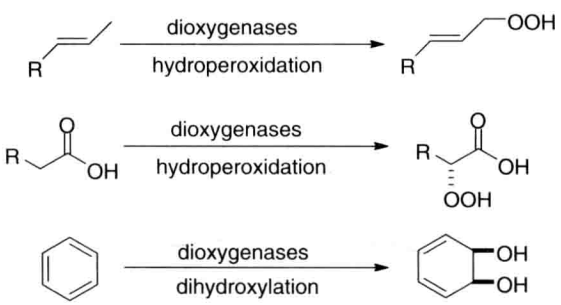


Figure 1.5

1.5.4 Alcohol Oxidases (EC 1.1.3.X)

Alcohol oxidases (AOs) catalyze oxidation of alcohols to aldehydes or ketones in the presence of molecular oxygen, with hydrogen peroxide being the usual by-product (Figure 1.6). Some of the most well-studied AOs are cholesterol oxidases, short-chain aliphatic alcohol oxidases, aromatic alcohol oxidases, pyranose oxidases, glycolate oxidases, glucose oxidases, galactose oxidases and nucleoside oxidases [38–40]. Cholesterol oxidases catalyze oxidation of allylic alcohol in the cholesterol scaffolds [41]. The regeneration of cofactor FAD is relatively easier, as it is tightly bound. Although these enzymes use oxygen, they can also be deactivated by oxygen or the hydrogen peroxide by-product. It is to be noted that peroxidases and cholorperoxidases can also catalyze the oxidation of alcohols using hydrogen peroxide (H_2O_2).