Enzymes in Industry

Production and Applications

Edited by Wolfgang Aehle

WILEY-VCH Verlag GmbH & Co. KGaA

Enzymes in Industry
Edited by W. Aehle

Also of Interest:

Liese, A., Seelbach, K., Wandrey, C.

Industrial Biotransformations

2000
3-527-30094-5

Dunn, I. J., Heinzle, E., Ingham, J., Přenosil, J. E.

Biological Reaction Engineeering
Dynamic Modelling and Simulation

Second, Completely Revised and Extended Edition

2003
ISBN 3-527-30759-1

Bisswanger, H.

Enzyme Kinetics
Principles and Methods

2002
ISBN 3-527-30343-X

Bisswanger, H.

Practical Enzymology

2003
ISBN 3-527-30444-4

Bornscheuer, U. T. (Ed.)

Enzymes in Lipid Modification

2000
ISBN 3-527-30176-3

Enzymes in Industry

Production and Applications

Edited by Wolfgang Aehle

WILEY-VCH Verlag GmbH & Co. KGaA

Dr. Wolfgang Aehle
Genencor International B. V.
PO Box 218
2300 AE Leiden
The Netherlands

First Edition 1990
Second, Completely Revised Edition 2004

Library of Congress Card No.: Applied for.
British Library Cataloguing-in-Publication Data:
A catalogue record for this book is available from the British Library

Die Deutsche Bibliothek –
CIP Catalo guing-in-Publication-Data
A catalogue record for this publication is available from Die Deutsche Bibliothek

Printed in the Federal Republic of Germany.

Printed on acid-free paper.

Typesetting Kühn & Weyh, Satz und Medien, Freiburg
Printing Strauss Offsetdruck GmbH, Mörlenbach
Bookbinding Litges & Dopf Buchbinderei GmbH, Heppenheim

ISBN 3-527-29592-5

Contents

List of Contributors

Editor

Dr. Wolfgang Aehle
Genencor International B. V.
Research & Development
P.O. Box 218
2300 AE Leiden
The Netherlands

Authors

Dr. Wolfgang Aehle
Genencor International B. V.
Research & Development
P.O. Box 218
2300 AE Leiden
The Netherlands
Chapter 1, Sections 4.2, 5.4

Dr. Richard L. Antrim
Grain Processing Corporation
1600 Oregon Street
Muscatine, IA 52761-1494
USA
Section 5.2.3

Todd Becker
Genencor International, Inc.
925 Page Mill Road
Palo Alto, CA 94304-1013
USA
Section 3.2.5

Dr. Rick Bott
Genencor International, Inc.
925 Page Mill Road
Palo Alto, CA 94304-1013
USA
Section 4.2

Dr. Johanna Buchert
VTT Biotechnology
P.O. Box 1500
2044 VTT
Finland
Section 5.2.6

Dr. Heidi Burrows
Finfeeds International
P.O. Box 777
Wiltshire, SN8 1XN Marlborough
UK
Section 5.2.4

Alice J. Caddow
Genencor International Inc.
925 Page Mill Road
Palo Alto, CA 94304-1013
USA
Chapter 7

Dr. Gopal K. Chotani
Genencor International, Inc.
925 Page Mill Road
Palo Alto, CA 94304-1013
USA
Section 3.1

Beth Concoby
Genencor International Inc.
925 Page Mill Road
Palo Alto, CA 94304-1013
USA
Chapter 7

Dr. Hans de Nobel
Genencor International B. V.
Research & Development
P.O. Box 218
2300 AE Leiden
The Netherlands
Section 4.1

Dr. André de Roos
DSM Food Specialities
P.O. Box 1
2600 MA Delft
The Netherlands
Section 5.1.4

Dr. Carlo Dinkel
SiChem GmbH
BITZ
Fahrenheitstr. 1
28359 Bremen
Germany
Section 6.1

Timothy C. Dodge
Genencor International, Inc.
925 Page Mill Road
Palo Alto, CA 94304-1013
USA
Section 3.1

Prof. Dr. Karlheinz Drauz
Degussa AG
Fine Chemicals
Rodenbacher Chaussée 4
63457 Hanau-Wolfgang
Germany
Section 6.1

Prof. Dr. Saburo Fukui †
formerly Department of
Industrial Chemistry
Faculty of Engineering
Kyoto University
606-8501 Kyoto
Japan
Section 3.3

Dr. Christian Gölker
Bayer AG
Aprather Weg
42096 Wuppertal
Germany
Section 3.2

Dr. Catherine Grassin
DSM Food Specialties
15, rue des Comtesses
P.O. Box 239
59472 Seclin Cedex
France
Section 5.1.2

Dr. Harald Gröger
Degussa AG
Project House Biotechnology
Rodenbacher Chaussée 4
63457 Hanau-Wolfgang
Germany
Section 6.1

Dr. Meng H. Heng
Genencor International, Inc.
925 Page Mill Road
Palo Alto, CA 94304-1013
USA
Sections 3.2.2.4, 3.2.4.1

Dr. Günther Henniger
formerly Roche Diagnostics GmbH
Nonnenwald 2
82377 Penzberg
Germany
Section 6.4

Dr. Ivan Herbots
Procter & Gamble Eurocor S.A.
Temselaan 100
1853 Strombeek-Bever
Belgium
Section 5.2.1

Dr. Marga Herweijer
DSM Food Specialties
Research and Development
P.O. Box 1
2600 MA Delft
The Netherlands
Section 5.1.2

Dr. Brian Jones
Genencor International B. V.
Research & Development
P.O. Box 218
2300 AE Leiden
The Netherlands
Section 4.1

Dr. Albert Jonke
Roche Diagnostics GmbH
Nonnenwald 2
82377 Penzberg
Germany
Chapter 2

John Kan
Genencor International, Inc.
925 Page Mill Road
Palo Alto, CA 94304-1013
USA
Sections 3.2.2.4, 3.2.4.1

Dr. Christoph Kessler
Roche Diagnostics GmbH
Nonnenwald 2
82377 Penzberg
Germany
Section 6.5

Dr. Beatrix Kottwitz
R&D / Technology Laundry and
Home Care
Henkel KGaA
Henkelstr. 67
40191 Düsseldorf
Germany
Section 5.2.2

Dr. Karsten M. Kragh
Danisco A/S
Edwin Rahrs Vej 38
8220 Brabrand
Denmark
Section 5.1.1

Dr. Georg-Burkhard Kresse
Roche Diagnostics GmbH
Pharma Research / Biology
Nonnenwald 2
82377 Penzberg
Germany
Section 6.2

Dr. Herman B. M. Lenting
TNO Industry
Centre for Textile Research
P.O. Box 337
7500 AH Enschede
The Netherlands
Section 5.2.5

Dr. Karl-Heinz Maurer
Henkel KGaA
Department of Enzyme Technology
Henkelstr. 67
40191 Düsseldorf
Germany
Section 5.3

Dr. Gerhard Michal
formerly Boehringer Mannheim GmbH
Research
68298 Mannheim
Germany
Chapter 2

Dr. Marja-Leena Niku-Paavola
VTT Biotechnology
P.O. Box 1500
2044 VTT
Finland
Section 5.2.6

Prof. Dr. Richard N. Perham
University of Cambridge
Department of Biochemistry
80 Tennis Court Road
CB2 1GA Cambridge
England
Chapter 1

Dr. Charlotte Horsmans Poulsen
Danisco A/S
Edwin Rahrs Vej 38
8220 Brabrand
Denmark
Section 5.1.1

Prof. Dr. Peter J. Reilly
Iowa State University
Department of Chemical Engineering
2114 Sweeney Hall
Ames, IA 50011-2230
USA
Section 5.2.3

Dr. Andrea Saettler
Henkel KGaA
Department of Enzyme Technology
Henkelstr. 67
40191 Düsseldorf
Germany
Section 5.3

Dr. Rainer Schmuck
Roche Diagnostics GmbH
Diagnostic Research
Nonnenwald 2
82377 Penzberg
Germany
Section 6.3

Dr. Carsten Schultz
EMBL
Meyerhofstr. 1
69117 Heidelberg
Germany
Section 6.1

Dr. Jorn Borch Soe
Danisco A/S
Edwin Rahrs Vej 38
8220 Brabrand
Denmark
Section 5.1.5

Jens Frisback Sorensen
Danisco A/S
Edwin Rahrs Vej 38
8220 Brabrand
Denmark
Section 5.1.1

Dr. Anna Suurnäkki
VTT Biotechnology
P.O. Box 1500
2044 VTT
Finland
Section 5.2.6

Prof. Dr. Atsuo Tanaka
Department of Industrial Chemistry
Faculty of Engineering
Kyoto University
606-8501 Kyoto
Japan
Section 3.3

Dr. Andreas Herman
Terwisscha van Scheltinga
DSM-Gist Research and Development
Alexander Fleminglaan 1
2613 AX Delft
The Netherlands
Section 3.1

Dr. Liisa Viikari
VTT Biotechnology
P.O. Box 1500
2044 VTT
Finland
Section 5.2.6

Prof. Dr. Herbert Waldmann
Max-Planck-Institut für molekulare
Physiologie
Otto-Hahn-Str. 11
44227 Dortmund
Germany
Section 6.1

Dr. Jan Wilms
Antonine der Kinderen
Rietschoot 179
1511 WG Oostzaan
The Netherlands
Section 5.1.3

Dr. Karl Wulff
Boeringer Mannheim GmbH
Bahnhofsstr. 9-15
82327 Tutzing
Germany
Section 6.3

Abbreviations

A:	adenosine
ACA:	acetamidocinnamic acid
ACL:	α-amino-ε-aprolactam
ADH:	alcohol dehydrogenase
ADI:	acceptable daily intake
ADP:	adenosine 5′-diphosphate
Ala:	alanine
Arg:	arginine
AMP:	adenosine 5′-monophosphate
ATC:	D,L-2-amino-D^2-thiazoline-4-carboxylic acid
ATP:	adenosine 5′-triphosphate
C:	cytidine
cDNA:	copy DNA
CL:	citrate lyase
CMP:	cytidine 5′-monophosphate
CoA:	coenzyme A
CS:	citrate synthetase
CTP:	cytidine 5′-triphosphate
d:	deoxy
dam:	gene locus for *E. coli* DNA adenine methylase (N^6-methyladenine)
dcmI:	gene locus for *E. coli* DNA cytosine methylase(5-methylcytosine)
dd:	dideoxy
ddNTP:	dideoxynucleoside 5′-triphosphate
DE:	dextrose equivalent
DEAE:	diethylaminoethyl
DNA:	deoxyribonucleic acid
DNase:	deoxyribonuclease
dNTP:	deoxynucleoside 5′-triphosphate
DOPA:	3-(3,4-dihydroxyphenylalanine) [3-hydroxy-L-tyrosine]
dpm:	decays per minute
ds:	double-stranded

E.C.: Enzyme Commission

F6P: fructose 6-phosphate
FAN: free alpha amino nitrogen, i.e., a measure of peptides/amino acids available for yeast to be used as nutrient
fMet: *N*-formylmethionine
FMN: flavin mononucleotide
$FMNH_2$: flavin mononucleotide, reduced

G: guanosine
GDP: guanosine 5′-diphosphate
Glu: glutamic acid
Gly: glycine
GMP: guanosine 5′-monophosphate
GOD: glucose oxidase
GOT: glutamate oxaloacetate transaminase
G6P: glucose 6-phosphate
GPT: glutamate pyruvate transaminase
GTP: guanosine 5′-triphosphate

3-HBDH: 3-hydroxybutyrate dehydrogenase
HFCS: high-fructose corn syrup
*hsd*M: *E. coli* gene locus for methylation
*hsd*R: *E. coli* gene locus for restriction
*hsd*S: *E. coli* gene locus for sequence specificity

IDP: inosine 5′-diphosphate
Ile: isoleucine
INT: iodonitrotetrazolium chloride
ITP: inosine 5′-triphosphate

LDH: lactate dehydrogenase
Lys: lysine

m(superscript): methylated
MDH: malate dehydrogenase
Met: methionine
M6P: mannose 6-phosphate
mRNA: messenger RNA
MTT: 3-(4,5-dimethylthiazolyl-2)-2,5-diphenyltetrazolium bromide

N: any nucleotide
NAD: nicotinamide – adenine dinucleotide
NADH: nicotinamide – adenine dinucleotide, reduced
NADP: nicotinamide – adenine dinucleotide phosphate

NADPH: nicotinamide – adenine dinucleotide phosphate, reduced
NMN: nicotinamide mononucleotide
NTP: nucleoside 5′-triphosphate

p: phosphate groups
^{32}P: phosphate groups containing ^{32}P phosphorus atoms
p_i: inorganic phosphate
°P: degree Plato; i.e., sugar content equivalent to 1 % sucrose by weight
PEP: phosphoenolpyruvate
6-PGDH: 6-phosphogluconate dehydrogenase
Phe: phenylalanine
PMS: 5-methylphenazinium methyl sulfate
poly(dA): poly(deoxyadenosine 5′-monophosphate)
ppi: inorganic pyrophosphate
Pro: proline
PRPP: phosphoribosyl pyrophosphate
Pu: purine
Py: pyrimidine

r: ribo
RNA: ribonucleic acid
RNase: ribonuclease

SAM: *S*-adenosylmethionine
SMHT: serine hydroxymethyltransferase
ss: single-stranded

T: thymidine
TMP: thymidine 5′-monophosphate
tRNA: transfer RNA
TTP: thymidine 5′-triphosphate

U: uridine
UMP: uridine 5′-monophosphate
UTP: uridine 5′-triphosphate

Val: valine

Plasmids:
pBR322
pBR328
pSM1
pSP64
pSP65
pSPT18, pSPT19

pT7–1, pT7–2
pUC 18, pUC 19
pUR222

Bacteriophages:
fd
ghl
M13
N4
PBS1
PBS2
SPO1
SP6
SP15
T3
T4
T5
T7
XP12
gt11
SM11
X174

Eukaryotic viruses:
Ad2
SV40

1
Introduction

Enzymes are the catalysts of biological processes. Like any other catalyst, an enzyme brings the reaction catalyzed to its equilibrium position more quickly than would occur otherwise; an enzyme cannot bring about a reaction with an unfavorable change in free energy unless that reaction can be coupled to one whose free energy change is more favorable. This situation is not uncommon in biological systems, but the true role of the enzymes involved should not be mistaken.

The activities of enzymes have been recognized for thousands of years; the fermentation of sugar to alcohol by yeast is among the earliest examples of a biotechnological process. However, only recently have the properties of enzymes been understood properly. Indeed, research on enzymes has now entered a new phase with the fusion of ideas from protein chemistry, molecular biophysics, and molecular biology. Full accounts of the chemistry of enzymes, their structure, kinetics, and technological potential can be found in many books and series devoted to these topics [1–5]. This chapter reviews some aspects of the history of enzymes, their nomenclature, their structure, and their relationship to recent developments in molecular biology.

1.1
History

Detailed histories of the study of enzymes can be found in the literature [6], [7].

Early Concepts of Enzymes. The term "enzyme" (literally "in yeast") was coined by Kühne in 1876. Yeast, because of the acknowledged importance of fermentation, was a popular subject of research. A major controversy at that time, associated most memorably with Liebig and Pasteur, was whether or not the process of fermentation was separable from the living cell. No belief in the necessity of vital forces, however, survived the demonstration by Buchner (1897) that alcoholic fermentation could by carried out by a cell-free yeast extract. The existence of extracellular enzymes had, for reasons of experimental accessibility, already been recognized. For example, as early as 1783, Spallanzani had demonstrated that gastric juice could digest meat in vitro, and Schwann (1836) called the active substance pepsin.

Kühne himself appears to have given trypsin its present name, although its existence in the intestine had been suspected since the early 1800s.

Enzymes as Proteins. By the early 1800s, the proteinaceous nature of enzymes had been recognized. Knowledge of the chemistry of proteins drew heavily on the improving techniques and concepts of organic chemistry in the second half of the 1800s; it culminated in the peptide theory of protein structure, usually credited to Fischer und Hofmeister. However, methods that had permitted the separation and synthesis of small peptides were unequal to the task of purifying enzymes. Indeed, there was no consensus that enzymes were proteins. Then, in 1926, Sumner crystallized urease from jack bean meal and announced it to be a simple protein. However, Willstätter argued that enzymes were not proteins but "colloidal carriers" with "active prosthetic groups". However, with the conclusive work by Northrop *et al.*, who isolated a series of crystalline proteolytic enzymes, beginning with pepsin in 1930, the proteinaceous nature of enzymes was established.

The isolation and characterization of intracellular enzymes was naturally more complicated and, once again, significant improvements were necessary in the separation techniques applicable to proteins before, in the late 1940s, any such enzyme became available in reasonable quantities. Because of the large amounts of accessible starting material and the historical importance of fermentation experiments, most of the first pure intracellular enzymes came from yeast and skeletal muscle. However, as purification methods were improved, the number of enzymes obtained in pure form increased tremendously and still continues to grow. Methods of protein purification are so sophisticated today that, with sufficient effort, any desired enzyme can probably be purified completely, even though very small amounts will be obtained if the source is poor.

Primary Structure. After the protein nature of enzymes had been accepted, the way was clear for more precise analysis of their composition and structure. Most amino acids had been identified by the early 20th century. The methods of amino acid analysis then available, such as gravimetric analysis or microbiological assay, were quite accurate but very slow and required large amounts of material. The breakthrough came with the work of Moore and Stein on ion-exchange chromatography of amino acids, which culminated in 1958 in the introduction of the first automated amino acid analyzer [8].

The more complex question – the arrangement of the constituent amino acids in a given protein, generally referred to as its primary structure – was solved in the late 1940s. The determination in 1951 of the amino acid sequence of the β-chain of insulin by Sanger and Tuppy [10] demonstrated for the first time that a given protein does indeed have a unique primary structure. The genetic implications of this were enormous. The introduction by Edman of the phenyl isothiocyanate degradation of proteins stepwise from the N-terminus, in manual form in 1950 and subsequently automated in 1967 [11], provided the principal chemical method for determining the amino acid sequences of proteins. The primary structures of pancreatic ribonuclease [12] and egg-white lysozyme [13] were published in 1963. Both of these

enzymes, simple extracellular proteins, contain about 120 amino acids. The first intracellular enzyme to have its primary structure determined was glyceraldehyde 3-phosphate dehydrogenase [14], which has an amino acid sequence of 330 residues and represents a size (250 – 400 residues) typical of many enzymes. Protein sequencing is increasingly performed by liquid chromatography/mass spectrometry (LC/MS) techniques, and several tools and software packages are now available for protein identification and characterization. The methods of protein sequence analysis are now so well developed that no real practical deterrent exists, other than time or expense, to determination of the amino acid sequence of any polypeptide chain [9].

A more recent fundamental concept called proteome (*prote*in complement to a gen*ome*) will enable researchers to unravel biochemical and physiological mechanisms of complex multivariate diseases at the functional molecular level. A new discipline, *proteomics*, complements physical genome research. Proteomics can be defined as "the qualitative and quantitative comparison of proteomes under different conditions to further unravel biological processes" [15].

Active Site. The fact that enzymes are highly substrate specific and are generally much larger than the substrates on which they act quickly became apparent. The earliest kinetic analyses of enzymatic reactions indicated the formation of transient enzyme – substrate complexes. These observations could be explained easily if the conversion of substrate to product was assumed to occur at a restricted site on an enzyme molecule. This site soon became known as the active center or, as is more common today, the active site.

Particular compounds were found to react with specific amino acid side chains and thus inhibit particular enzymes. This suggested that such side chains might take part in the catalytic mechanisms of these enzymes. An early example was the inhibition of glycolysis or fermentation by iodoacetic acid, which was later recognized as resulting from reaction with a unique cysteine residue of glyceraldehyde 3-phosphate dehydrogenase, which normally carries the substrate in a thioester linkage [16].

Many such group-specific reagents have now been identified as inhibitors of individual enzymes; often they are effective because of the hyper-reactivity of a functionally important side chain in the enzyme's active site. However, a more sophisticated approach to the design of enzyme inhibitors became possible when the reactive group was attached to a substrate; in this way, the specificity of the target enzyme was utilized to achieve selective inhibition of the enzyme [17]. Such active-site-directed inhibitors have acquired major importance not only academically in the study of enzyme mechanisms but also commercially in the search for a rational approach to selective toxicity or chemotherapy.

Three-Dimensional Structure. Chemical studies showed that the active site of an enzyme consists of a constellation of amino acid side chains brought together spatially from different parts of the polypeptide chain. If this three-dimensional structure was disrupted by denaturation, that is, without breaking any covalent bonds, the biological activity of the enzyme was destroyed. In addition, it was found that all

the information required for a protein to fold up spontaneously in solution and reproduce its native shape was contained in its primary structure. This was part of the original "central dogma" of molecular biology.

The X-ray crystallography of proteins [18] demonstrated unequivocally that a given protein has a unique three-dimensional structure. Among the basic design principles was the tendency of hydrophobic amino acid side chains to be associated with the hydrophobic interior of the folded molecule, whereas charged side chains were almost exclusively situated on the hydrophilic exterior or surface. The first high-resolution crystallographic analysis of an enzyme, egg-white lysozyme, confirmed these principles and led to the proposal of a detailed mechanism [19]. The active site was located in a cleft in the structure (Fig. 1), which has subsequently proved to be a common feature of active sites. According to this, the enzymatic reaction takes place in a hydrophobic environment, and the successive chemical events involving substrate and protein side chains are not constrained by the ambient conditions of aqueous solution and neutral pH.

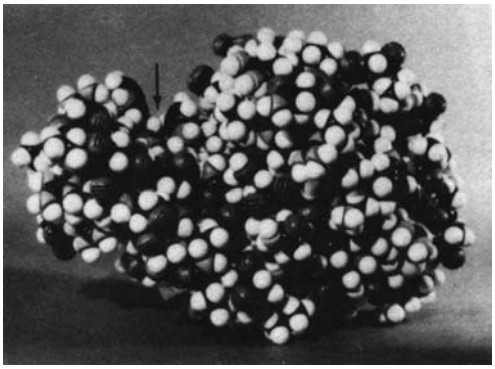

Figure 1. A molecular model of the enzyme lysozyme: the arrow points to the cleft that accepts the polysaccharide substrate (Reproduced by courtesy of J. A. Rupley)

1.2 Enzyme Nomenclature

Strict specificity is a distinguishing feature of enzymes, as opposed to other known catalysts. Enzymes occur in myriad forms and catalyze an enormous range of reactions. By the late 1950 s the number of known enzymes had increased so rapidly that their nomenclature was becoming confused or, worse still, misleading because the same enzyme was often known to different workers by different names; in addition, the name frequently conveyed little or nothing about the nature of the reaction catalyzed.

To bring order to this chaotic situation, an International Commission on Enzymes was established in 1956 under the auspices of the International Union of Biochem-